



## Reconstructing human pancreatic differentiation by mapping specific cell populations during development

Ramond, Cyrille; Glaser, Nicolas; Berthault, Claire; Ameri, Jacqueline; Kirkegaard, Jeannette Schlichting; Hansson, Mattias; Honoré, Christian; Semb, Tor Henrik; Scharfmann, Raphaël

*Published in:*  
eLife

*DOI:*  
[10.7554/eLife.27564](https://doi.org/10.7554/eLife.27564)

*Publication date:*  
2017

*Document version*  
Publisher's PDF, also known as Version of record

*Citation for published version (APA):*  
Ramond, C., Glaser, N., Berthault, C., Ameri, J., Kirkegaard, J. S., Hansson, M., Honoré, C., Semb, T. H., & Scharfmann, R. (2017). Reconstructing human pancreatic differentiation by mapping specific cell populations during development. *eLife*, 6, [e27564]. <https://doi.org/10.7554/eLife.27564>

# Reconstructing human pancreatic differentiation by mapping specific cell populations during development

Cyrille Ramond<sup>1,2,3</sup>, Nicolas Glaser<sup>1,2,3</sup>, Claire Berthault<sup>4</sup>, Jacqueline Ameri<sup>5</sup>, Jeannette Schlichting Kirkegaard<sup>6</sup>, Mattias Hansson<sup>7</sup>, Christian Honoré<sup>6</sup>, Henrik Semb<sup>5</sup>, Raphaël Scharfmann<sup>1,2,3\*</sup>

<sup>1</sup>INSERM U1016, Cochin Institute, Paris, France; <sup>2</sup>CNRS UMR 8104, Paris, France; <sup>3</sup>University of Paris Descartes, Paris, France; <sup>4</sup>Immunology Department, Unit for Lymphopoiesis, Paris, France; <sup>5</sup>The Danish Stem Cell Center (DanStem), Faculty of Health Sciences, University of Copenhagen, Denmark, Europe; <sup>6</sup>Department of Islet and Stem Cell Biology, Novo Nordisk A/S, Denmark, Europe; <sup>7</sup>Global Research External Affairs, Novo Nordisk A/S, Denmark, Europe

**Abstract** Information remains scarce on human development compared to animal models. Here, we reconstructed human fetal pancreatic differentiation using cell surface markers. We demonstrate that at 7 weeks of development, the glycoprotein 2 (GP2) marks a multipotent cell population that will differentiate into the acinar, ductal or endocrine lineages. Development towards the acinar lineage is paralleled by an increase in GP2 expression. Conversely, a subset of the GP2<sup>+</sup> population undergoes endocrine differentiation by down-regulating GP2 and CD142 and turning on *NEUROG3*, a marker of endocrine differentiation. Endocrine maturation progresses by up-regulating *SUSD2* and lowering *ECAD* levels. Finally, in vitro differentiation of pancreatic endocrine cells derived from human pluripotent stem cells mimics key in vivo events. Our work paves the way to extend our understanding of the origin of mature human pancreatic cell types and how such lineage decisions are regulated.

DOI: [10.7554/eLife.27564.001](https://doi.org/10.7554/eLife.27564.001)

\*For correspondence: raphael.scharfmann@inserm.fr

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 18

**Received:** 07 April 2017

**Accepted:** 17 July 2017

**Published:** 21 July 2017

**Reviewing editor:** Neil A Hanley, University of Manchester, United Kingdom

© Copyright Ramond et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Introduction

Intensive efforts are currently dedicated towards the development of cell replacement therapies using cell types derived from human pluripotent stem cells (hPSC). Human insulin-producing beta cells represent a paradigm for this type of objective. These cells have a major physiological function, regulating circulating glucose levels by producing and secreting insulin. In patients suffering from type one diabetes, these cells are destroyed by an autoimmune mechanism, and would thus need to be replaced (*Benthuyzen et al., 2016*). Beta cell replacement holds immense promises for diabetic patients and current strategies have reached major milestones (*Pagliuca et al., 2014; Russ et al., 2015; Rezanian et al., 2014*). However, it is well accepted that a more detailed understanding of beta cell development in human is required to generate unlimited functional human beta cells (*Johnson, 2016*).

The adult pancreas is composed of acinar cells that excrete enzymes into the duodenum through a ductal tree, and of endocrine cells (approximately 1% of the total pancreatic cells) that are clustered together forming the islets of Langerhans. The endocrine cells secrete hormones such as insulin (beta cells), glucagon (alpha cells) somatostatin (delta cells), pancreatic polypeptide (gamma cells) and ghrelin (epsilon cells). The pancreas develops from the primitive gut tube that evaginates into a dorsal and a ventral anlage (*Pan and Wright, 2011; Jennings et al., 2015*). Multipotent epithelial

pancreatic progenitors co-expressing the transcription factors PDX1 and NKX6-1 (Nelson et al., 2007; Cebola et al., 2015) proliferate upon signals (such as FGF10) from the adjacent mesenchyme (Bhushan et al., 2001) and subsequently differentiate into the acinar, ductal and endocrine lineages. Endocrine commitment is initially marked by the expression of a basic helix-loop-helix transcription factor, NEUROG3, and followed by the expression of the mature endocrine markers (Gradwohl et al., 2000; Gu et al., 2002).

Due to the difficulties associated with procuring staged human fetal tissues and the limited tools for their analysis, few data on human development is available and the majority of knowledge on tissue development derives from animal models. This also applies to the pancreas (Jennings et al., 2015) where only a limited number of studies have been performed on human pancreatic development. These studies demonstrate similarities but also differences between rodent and human pancreatic development (Scharfmann et al., 2013; Jennings et al., 2015; Nair and Hebrok, 2015).

Knowledge on human pancreatic development remains limited. More information exists concerning human hematopoietic cell differentiation thanks to the characterization and use of cell surface antigens that enabled to identify, quantify and purify hematopoietic stem cells and progenitors at different stages of their development (Eaves, 2015). By mirroring the hematopoietic field, we developed here an approach where cell surface markers are used to recapitulate the hierarchical sequence of human pancreatic development. Specifically, we characterized the expression levels of specific markers at different stages of human pancreatic development corresponding to 7 to 12 weeks of development (WD). First we purified human pancreatic epithelial cells by selecting cells positive for the transmembrane glycoprotein EPCAM and by excluding CD45<sup>+</sup> hematopoietic and CD31<sup>+</sup> endothelial cells. Next we segregated pancreatic epithelial cells into four populations based on the GP2 and CDH1 (ECAD) expression levels. We observed that the expression levels of GP2 and ECAD correlate with acinar, ductal and endocrine functions. By using the additional cell surface markers CD142 and SUSP2 we further refined endocrine cell differentiation. Finally, our development model also applies to the in vitro differentiation of hPSCs into pancreatic endocrine cells.

Taken together our work provides a novel approach to study human fetal pancreas development and bridges the path between in vivo and in vitro differentiation of human pancreatic endocrine cells.

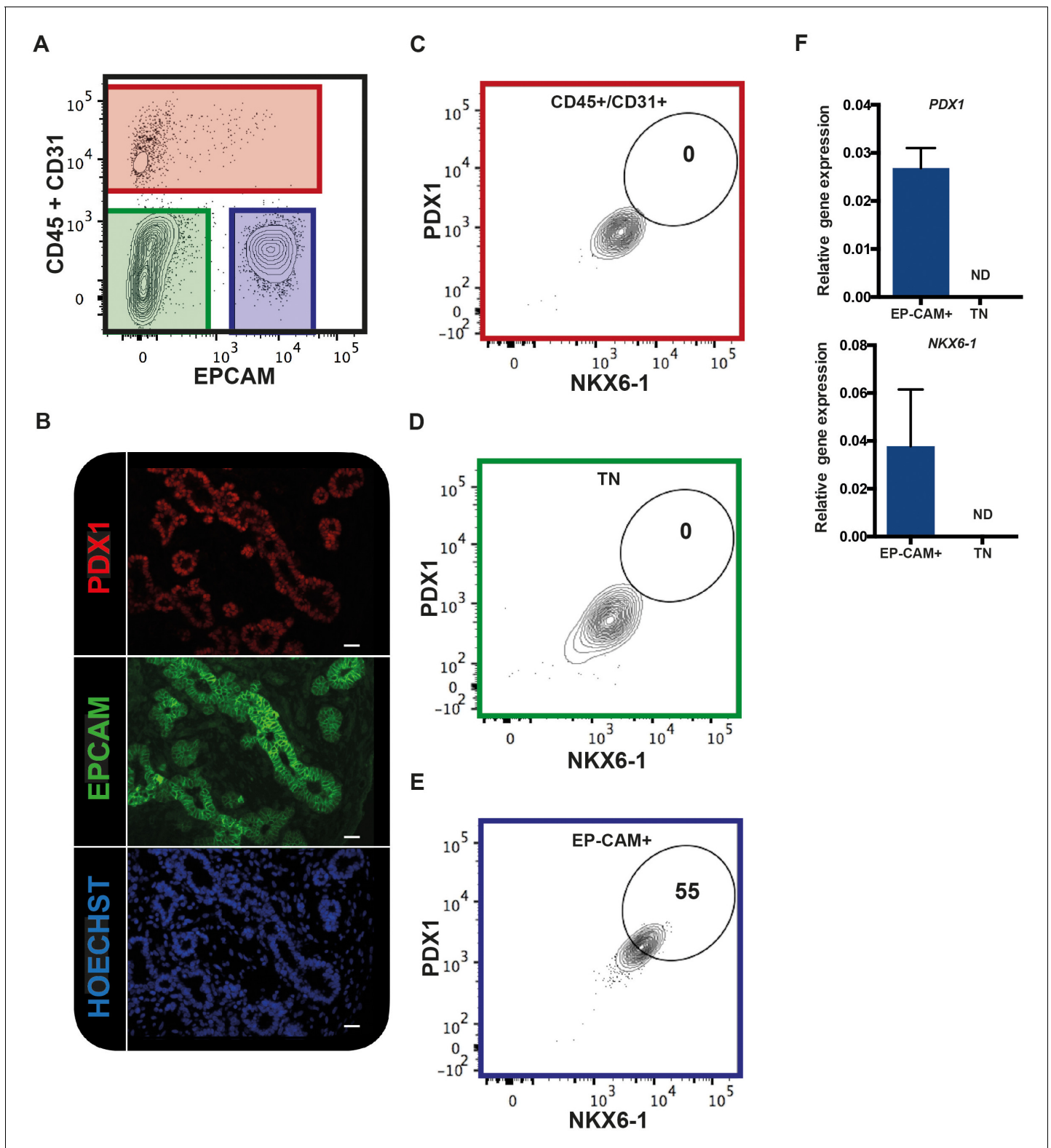
## Results

### EPCAM expression is restricted to the epithelium in the human fetal pancreas

We tested if EPCAM can be used to purify human fetal pancreatic epithelial fraction enriched in pancreatic progenitors. CD31 and CD45 were used to exclude endothelial and hematopoietic cells respectively. Using this combination of antibodies on human fetal pancreatic cells (9.7WD), we detected three distinct fractions: the CD45<sup>+</sup>/CD31<sup>+</sup> fraction (the endothelial/hematopoietic cells), the CD45<sup>-</sup>CD31<sup>-</sup>EPCAM<sup>-</sup> fraction, and the CD45<sup>-</sup>CD31<sup>-</sup>EPCAM<sup>+</sup> fraction (Figure 1A). To unveil which fraction contained the pancreatic progenitors, we assayed the expression of PDX1 and NKX6-1. Immunohistochemistry analysis of human fetal pancreatic sections showed that EPCAM<sup>+</sup> cells expressed PDX1 (Figure 1B). FACS analysis demonstrated that PDX1 and NKX6-1 were co-expressed in the CD45<sup>-</sup>CD31<sup>-</sup>EPCAM<sup>+</sup> fraction (Figure 1C–E). RT-qPCR analysis on sorted fractions confirmed that PDX1 and NKX6-1 expressions were restricted to the CD45<sup>-</sup>CD31<sup>-</sup>EPCAM<sup>+</sup> fraction (Figure 1F) whereas the CD45<sup>-</sup>CD31<sup>-</sup>EPCAM<sup>-</sup> fraction did not express pancreatic markers and most likely represent the mesenchymal pancreatic fraction (later referred as population M) (Figure 1A). These results suggest that a combination of the cell surface markers EPCAM, CD45, and CD31 can be used to purify the human fetal pancreatic epithelial fraction.

### GP2 and ECAD define four populations in the human fetal pancreatic epithelium that develop sequentially

Recently, GP2 was identified as a novel cell surface marker of the immature pancreatic progenitor cells derived from hPSC (Ameri et al., 2017). Furthermore, additional data indicate that ECAD levels are tightly modulated during endocrine differentiation (Gouzi et al., 2011). Therefore, we analyzed GP2 and ECAD expression levels in the CD45<sup>-</sup>CD31<sup>-</sup>EPCAM<sup>+</sup> fraction at 9.4WD. GP2 expression in



**Figure 1.** EPCAM expression in the human fetal pancreas. (A) The flow cytometry plot represents CD45 and CD31 expression against EPCAM gated on live human fetal pancreatic cells (9.7WD),  $n = 9$ . (B) Immunohistochemistry for PDX1 and EPCAM on pancreatic section (9WD),  $n = 3$ . Scale bar = 100  $\mu\text{m}$ . (C–E) Flow cytometry plots of PDX1 and NKX6-1 expression at 9.4WD on CD45<sup>+</sup>/CD31<sup>+</sup> cells (red square), CD45<sup>+</sup>CD31<sup>+</sup>EPCAM<sup>+</sup> cells (TN = triple negative green square) and CD45<sup>+</sup>CD31<sup>+</sup>EPCAM<sup>+</sup> cells (blue square). (F) RT-qPCR analysis of *PDX1* and *NKX6-1* expression on sorted CD45<sup>+</sup>CD31<sup>+</sup>EPCAM<sup>+</sup> and TN cells. ND = Not Detected.

DOI: [10.7554/eLife.27564.002](https://doi.org/10.7554/eLife.27564.002)



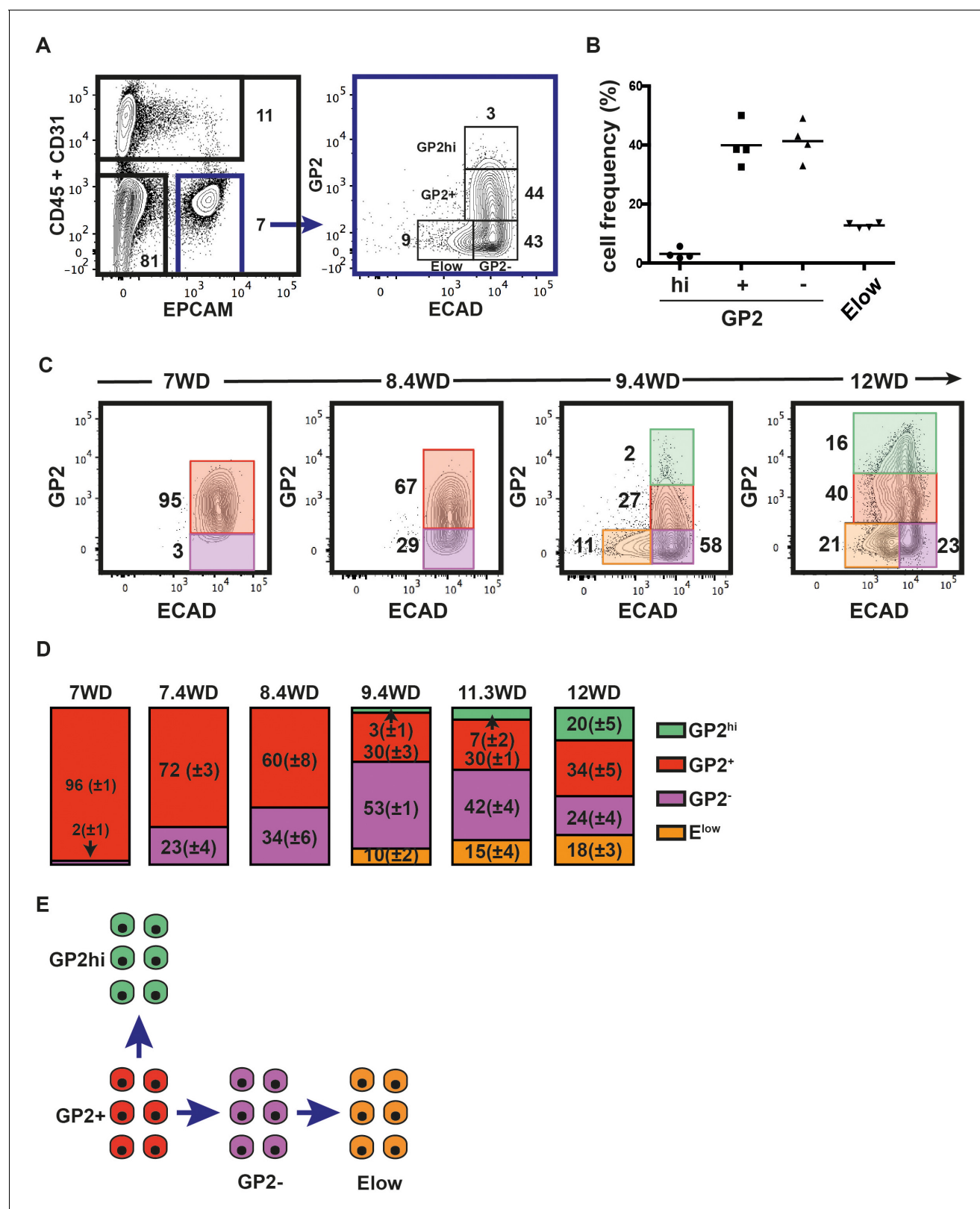
CD45<sup>+</sup>CD31<sup>+</sup> was restricted to the EPCAM<sup>+</sup> fraction (**Figure 2—figure supplement 1**). GP2 and ECAD expression segregated the CD45<sup>+</sup>CD31<sup>+</sup>EPCAM<sup>+</sup> fraction into four distinct populations: GP2<sup>hi</sup>ECAD<sup>+</sup> (named GP2<sup>hi</sup>), GP2<sup>+</sup>ECAD<sup>+</sup> (named GP2<sup>+</sup>), GP2<sup>+</sup>ECAD<sup>+</sup> (named GP2<sup>+</sup>), and the GP2<sup>+</sup>ECAD<sup>low</sup> (named E<sup>low</sup>) (**Figure 2A**). Remarkably, the cell frequencies of the four populations were well conserved from one pancreas to the other, demonstrating inter-individual homogeneity (**Figure 2B**). We assayed the dynamic expression of GP2 and ECAD in the pancreatic epithelial fraction (CD45<sup>+</sup>CD31<sup>+</sup>EPCAM<sup>+</sup>) during development (from 7WD to 12WD) (**Figure 2C**). At 7WD the epithelial fraction was essentially GP2<sup>+</sup> (96 ± 1%). From 7WD to 8.4WD, the GP2<sup>+</sup> population evolves from 2 ± 1% to 34 ± 6% (p<0.05) (**Figure 2C,D**). Interestingly, from 8.6WD, we detected cells with a lower ECAD level (**Figure 2—figure supplement 2A,B**). This E<sup>low</sup> population further increased in frequency from 10 ± 2% at 9.4WD to 18 ± 3% at 12WD (p<0.05) (**Figure 2C,D**). At 9.4WD, we also detected a GP2<sup>hi</sup> population, increasing in frequency during development (3 ± 1% at 9.4WD to 20 ± 5% at 12WD; p<0.05) (**Figure 2C,D** and **Figure 2—figure supplement 2C**). Altogether, the temporal and continuum GP2 and ECAD expression suggest a progeny relationship between specific cell populations, the GP2<sup>+</sup> population would differentiate either into a GP2<sup>hi</sup> population or into a GP2<sup>+</sup> population that would later decrease its ECAD level to give rise to the E<sup>low</sup> population (**Figure 2E**).

### Acinar and endocrine functions segregate within the GP2 and ECAD populations

To characterize the four epithelial populations described above, we sorted the GP2<sup>hi</sup>, GP2<sup>+</sup>, GP2<sup>+</sup> and E<sup>low</sup> populations and performed global transcriptomic analyses combined with RT-qPCR analyses at 9 and 11WD. As a non-epithelial control, we included the CD45<sup>+</sup>CD31<sup>+</sup>EPCAM<sup>+</sup> fraction (population M) (**Figure 1A**). Due to limitations in cell numbers, the GP2<sup>hi</sup> population was only sorted at 11WD. Principal component analysis (PCA) on the sorted populations at 9WD revealed three clusters. PC1 separated the epithelial populations (GP2<sup>hi</sup>, GP2<sup>+</sup>, GP2<sup>+</sup> and E<sup>low</sup>) from the mesenchymal fraction (M), while PC2 segregated the E<sup>low</sup> population from the GP2<sup>hi</sup>, GP2<sup>+</sup> and GP2<sup>+</sup> populations (**Figure 3A**). Gene Set Enrichment Analysis (GSEA) using Gene Ontology database indicated that digestion was the most represented biological process in the GP2<sup>hi</sup> and GP2<sup>+</sup> populations while endocrine functions (insulin and peptide secretion, hormone secretion) were enriched in the E<sup>low</sup> population (**Figure 3B**, **Figure 3—figure supplement 2**). Conversely the pancreatic functions appeared less enriched in the GP2<sup>+</sup> population. Next, we defined the list of the 'specific enriched genes' per population (**Supplementary file 1b–c** and Materials and method section). We compared the fetal 'specific enriched genes' with the RNAseq Single Cell data from human adult pancreas (**Segerstolpe et al., 2016**) (**Figure 3—figure supplement 3**). The GP2<sup>hi</sup> population displayed 27 enriched genes (p<0.05), 85% being preferentially expressed in the adult acinar cells (**Figure 3C**, and **Figure 3—figure supplement 3A**). In contrast, the E<sup>low</sup> population contained 91 (at 9WD) and 34 (at 11WD) differentially expressed genes (p<0.05) that were also enriched (98% and 100% respectively) in the adult endocrine cells (alpha, beta, delta, epsilon or gamma cells) (**Figure 3C**, and **Figure 3—figure supplement 3B**).

We next generated heatmaps based on Gene Ontology lists and selected acinar, ductal and endocrine genes. By RT-qPCR analyses we confirmed that acinar markers such as *CEL*, *CELA3A* and *CTRC* were enriched in the GP2<sup>+</sup> population at 9WD and in the GP2<sup>hi</sup> population at 11WD (**Figure 3—figure supplement 2A**, **Figure 4A**). Ductal markers (*CFTR*, *KRT19* and *SPP1*) were enriched in the GP2<sup>+</sup> and GP2<sup>+</sup> populations, especially *CFTR* at 9WD and 11–13WD (**Figure 4—figure supplement 1**). The E<sup>low</sup> populations (9 and 11WD) were enriched with endocrine transcription factors such as *ARX*, *PAX4*, *PAX6*, *NEUROD1*, *MAFA*, *MAFB*, *ISL1* and *NKX2.2* and hormones like *CHGA*, *GCG*, *GHRL* and *INS* (**Figure 3—figure supplement 2B**, **Figure 4B**).

Finally, to define the cell population where endocrine cells first differentiate, we followed the expression of the endocrine progenitor marker *NEUROG3*. *NEUROG3* was first detected at 8.4WD in the GP2<sup>+</sup> population (**Figure 4C**) prior to the detection of the E<sup>low</sup> population (**Figure 2C**) and next (9 to 13WD), found enriched in the E<sup>low</sup> population (**Figure 4C**). Thus, our data suggest that the first human pancreatic endocrine progenitors differentiate in the GP2<sup>+</sup> population and mature while decreasing ECAD levels in the E<sup>low</sup> population.



**Figure 2.** GP2 and ECAD expression in the human fetal pancreatic epithelium. GP2 and ECAD expressions were assayed by flow cytometry during development. (A) FACS plots display the expression at 9.4WD of CD45 and CD31 against EPCAM (left plot) and GP2 and ECAD gated on CD45<sup>+</sup>CD31<sup>+</sup>EPCAM<sup>+</sup> (right plot).  $n = 4$  (B) Cell frequencies of the GP2<sup>hi</sup> (GP2<sup>hi</sup>ECAD<sup>+</sup>), GP2<sup>+</sup> (GP2<sup>+</sup>ECAD<sup>+</sup>), GP2<sup>-</sup> (GP2<sup>-</sup>ECAD<sup>+</sup>) and E<sup>low</sup> (GP2<sup>-</sup>ECAD<sup>low</sup>) populations at 9.4WD.  $n = 4$  (mean  $\pm$  SEM) (C) GP2 and ECAD expressions on fetal pancreases at 7-12WD gated on CD45<sup>+</sup>CD31<sup>+</sup>EPCAM<sup>+</sup> cells. 7WD Figure 2 continued on next page

Figure 2 continued

n = 2, 8.4WD n = 9, 9.4WD n = 4, 12WD n = 5. (mean  $\pm$  SEM) (D) Cell frequencies of the GP2<sup>hi</sup>, GP2<sup>+</sup>, GP2<sup>-</sup> and E<sup>low</sup> populations at 7-12WD. Cell frequencies were calculated from three independent experiments for each time point. (E) Scheme that represents the development of GP2<sup>hi</sup>, GP2<sup>+</sup>, GP2<sup>-</sup> and E<sup>low</sup> populations.

DOI: [10.7554/eLife.27564.003](https://doi.org/10.7554/eLife.27564.003)

The following source data and figure supplements are available for figure 2:

**Source data 1.** Cell frequency at 9.4WD by flow cytometry.

DOI: [10.7554/eLife.27564.004](https://doi.org/10.7554/eLife.27564.004)

**Source data 2.** Cell frequency during development.

DOI: [10.7554/eLife.27564.005](https://doi.org/10.7554/eLife.27564.005)

**Figure supplement 1.** Gating strategy for GP2 and ECAD.

DOI: [10.7554/eLife.27564.006](https://doi.org/10.7554/eLife.27564.006)

**Figure supplement 2.** Expression of ECAD and GP2 during development.

DOI: [10.7554/eLife.27564.007](https://doi.org/10.7554/eLife.27564.007)

## CD142 and SUSD2 reveal heterogeneity within the GP2<sup>-</sup> and E<sup>low</sup> populations during development

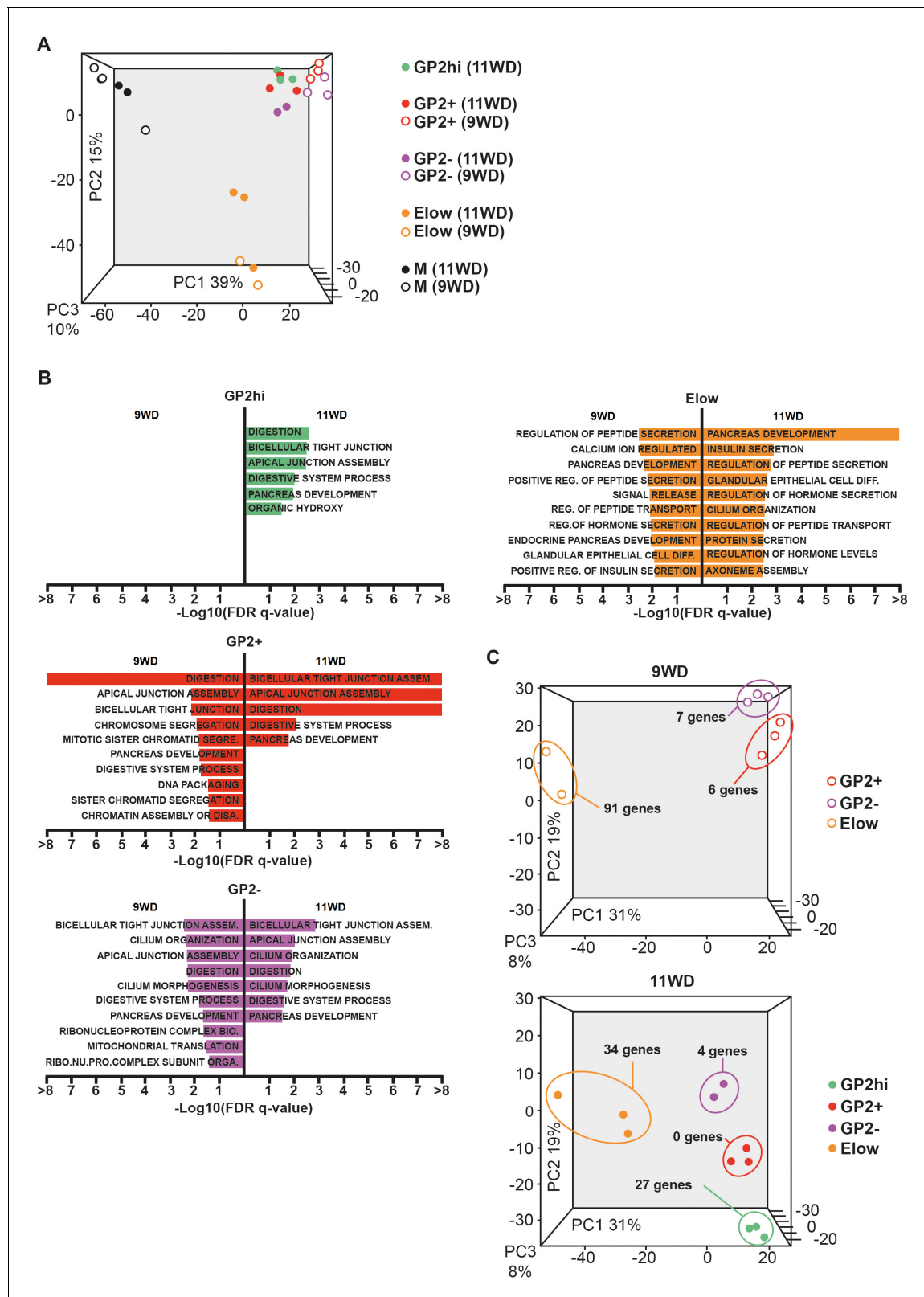
The GP2<sup>-</sup> population displayed both duct and endocrine progenitor cells markers suggesting it contains more than one cell type. We therefore sought for additional discriminant markers by scrutinizing our transcriptomic data. We observed two cell surface markers, CD142 and SUSD2 with opposite expression patterns: the E<sup>low</sup> population expressed lower CD142 and higher SUSD2 mRNA levels than the GP2<sup>+</sup> and GP2<sup>-</sup> populations (**Figure 5A**). FACS analyses showed that at 9.4WD, the GP2<sup>hi</sup> and GP2<sup>+</sup> populations were uniformly CD142<sup>+</sup>SUSD2<sup>-</sup>, while the GP2<sup>-</sup> and E<sup>low</sup> populations were further divided into three subsets: CD142<sup>+</sup>SUSD2<sup>-</sup>, CD142<sup>-</sup>SUSD2<sup>-</sup> and CD142<sup>-</sup>SUSD2<sup>+</sup> (**Figure 5B**). Noteworthy, the CD142<sup>-</sup>SUSD2<sup>+</sup> subset was scarce in the GP2<sup>-</sup> population (6%), but represented 40% of the E<sup>low</sup> population at 9.4WD (**Figure 5B**). Accordingly, our data reflects the heterogeneity within the GP2<sup>-</sup> and E<sup>low</sup> populations that is further resolved using CD142 and SUSD2 antibodies.

Next we examined the CD142 and SUSD2 expression patterns in the GP2<sup>-</sup> and in the E<sup>low</sup> populations during development. At 7WD the majority of the GP2<sup>-</sup> population was CD142<sup>+</sup>SUSD2<sup>-</sup> (95  $\pm$  2%). This frequency gradually decreased to 62  $\pm$  2% at 9.4WD while the frequency of the CD142<sup>-</sup>SUSD2<sup>-</sup> subset increased from 5  $\pm$  2% at 7WD to 43% at 11.3WD  $p < 0.05$  (**Figure 5C,D**). We detected the first SUSD2<sup>+</sup> cells in the GP2<sup>-</sup>CD142<sup>-</sup> subset at a low frequency (1%) at 8.4WD (**Figure 5C**). The first E<sup>low</sup> population was detected at 8.6WD and was divided into CD142<sup>+</sup>SUSD2<sup>-</sup>, CD142<sup>-</sup>SUSD2<sup>-</sup> and CD142<sup>-</sup>SUSD2<sup>+</sup> subsets (**Figure 2—figure supplement 2A,B** and **Figure 5—figure supplement 1**). As development progressed the frequency of the E<sup>low</sup>CD142<sup>-</sup>SUSD2<sup>+</sup> subset decreased (from 48  $\pm$  7% at 8.6WD to 10  $\pm$  1% at 12WD;  $p < 0.05$ ), while the frequency of the E<sup>low</sup>CD142<sup>-</sup>SUSD2<sup>-</sup> subset increased (from 30  $\pm$  4% at 8.6WD to 79  $\pm$  3% at 12WD;  $p < 0.05$ ) (**Figure 5B,C,E**). Our data thus indicate that the first GP2<sup>-</sup> cells are CD142<sup>+</sup>. They progress in their differentiation program by down-regulating CD142, then up-regulating SUSD2 and finally decreasing ECAD levels.

## Endocrine progenitors develop in the GP2<sup>-</sup>CD142<sup>-</sup>SUSD2<sup>-</sup> subset and mature within the E<sup>low</sup>SUSD2<sup>+</sup> subset

We analyzed the expression pattern of endocrine markers from 8.6WD to 13WD in the three E<sup>low</sup> subsets. The E<sup>low</sup>CD142<sup>+</sup>SUSD2<sup>-</sup> subset did not express endocrine markers such as CHGA and NEUROG3 (**Figure 6—figure supplement 1**). At 8.6WD, CHGA, NEUROD1, NKX2-2 and INS were solely detected in the E<sup>low</sup>CD142<sup>-</sup>SUSD2<sup>+</sup> subset (**Figure 6A**). Interestingly, at 10-12WD, CHGA, NEUROD1 and NKX2-2 expressions were detected both in the E<sup>low</sup>CD142<sup>-</sup>SUSD2<sup>+</sup> and the E<sup>low</sup>CD142<sup>-</sup>SUSD2<sup>-</sup> subsets while INS was exclusively detected in the E<sup>low</sup>CD142<sup>-</sup>SUSD2<sup>-</sup> subset (**Figure 6B**).

Next we assessed NEUROG3 expression in the different subsets. The first NEUROG3<sup>+</sup> were detected in the GP2<sup>-</sup>CD142<sup>-</sup> subset at 8.4WD before SUSD2 up-regulation (**Figure 6C**). From 8.6WD, when the E<sup>low</sup> population was first detected NEUROG3 was enriched in the E<sup>low</sup>CD142<sup>-</sup>SUSD2<sup>+</sup> subset (**Figure 6C**). Our results indicate that endocrine progenitors first appear



**Figure 3.** Transcriptomic analysis of the GP2<sup>hi</sup>, GP2<sup>+</sup>, GP2<sup>-</sup> and E<sup>low</sup> populations. (A) PCA map of sorted pancreatic cells (epithelium and mesenchyme). (B) Top enriched biological processes in each cell population compared to the mesenchyme. Results were obtained with GSEA software using the GO database. (C) PCA map of epithelial-sorted cells (GP2<sup>hi</sup>, GP2<sup>+</sup>, GP2<sup>-</sup> and E<sup>low</sup> populations) from 9 (top map) and 11WD (bottom map). The number of

Figure 3 continued on next page

Figure 3 continued

specifically genes enriched in each population in each population ( $p < 0.05$ ) is displayed. PCA maps are displayed in 2D with the three Principal Components on figure **Figure 3—figure supplement 1**) and the Gene Ontology lists in **Supplementary file 1a**.

DOI: [10.7554/eLife.27564.008](https://doi.org/10.7554/eLife.27564.008)

The following figure supplements are available for figure 3:

**Figure supplement 1.** Principal component analysis.

DOI: [10.7554/eLife.27564.009](https://doi.org/10.7554/eLife.27564.009)

**Figure supplement 2.** Heatmaps of genes from GO Biological processes digestion, insulin secretion and pancreas development.

DOI: [10.7554/eLife.27564.010](https://doi.org/10.7554/eLife.27564.010)

**Figure supplement 3.** The GP2<sup>hi</sup> and E<sup>low</sup> populations are enriched in acinar and endocrine markers respectively.

DOI: [10.7554/eLife.27564.011](https://doi.org/10.7554/eLife.27564.011)

in the GP2<sup>CD142</sup><sup>SUSD2</sup><sup>-</sup> subset and next mature within the E<sup>low</sup>CD142<sup>SUSD2</sup><sup>+</sup> subset. *INS* is first detected in the E<sup>low</sup>CD142<sup>SUSD2</sup><sup>+</sup> subset and later on in the E<sup>low</sup>CD142<sup>SUSD2</sup><sup>-</sup> subset.

### The E<sup>low</sup>CD142<sup>SUSD2</sup><sup>+</sup> subset predominantly contains endocrine progenitors

To assess the frequency of cells expressing endocrine progenitor (*NEUROG3*, *NEUROD1*, *NKX2-2*) and duct (*CFTR*) markers in the GP2<sup>CD142</sup><sup>SUSD2</sup><sup>-</sup> and in the E<sup>low</sup>CD142<sup>SUSD2</sup><sup>+</sup> subsets, we performed single cell RT-qPCR analyses at 9WD. The GP2<sup>CD142</sup><sup>SUSD2</sup><sup>-</sup> subset was heterogeneous and composed of 78% *CFTR*<sup>+</sup>*NEUROG3* cells, 16% *NEUROG3*<sup>+</sup> and 6% *CFTR*<sup>+</sup>*NEUROG3*<sup>-</sup> cells. Among the *NEUROG3*<sup>+</sup>, 47% were *NEUROD1*<sup>NKX2.2</sup>, 43% co-expressed *NEUROD1*, and 10% co-expressed *NEUROD1* and *NKX2.2*. Interestingly, 60% of *NEUROG3*<sup>+</sup> were *CFTR*<sup>+</sup> (**Figure 6D**). Conversely, only 2% of the E<sup>low</sup>CD142<sup>SUSD2</sup><sup>+</sup> subset was *CFTR*<sup>+</sup>*NEUROG3*<sup>-</sup>, while 74% was *NEUROG3*<sup>+</sup> and 24% *CFTR*<sup>+</sup>*NEUROG3*<sup>-</sup>. *NEUROG3*<sup>+</sup> were essentially homogenous as 100% expressed *NEUROD1*<sup>+</sup> and 83% co-expressed *NKX2-2* as well (**Figure 6E**). Our findings demonstrate that the majority of the E<sup>low</sup>CD142<sup>SUSD2</sup><sup>+</sup> subset is *NEUROG3*<sup>+</sup> co-expressing *NEUROD1* and *NKX2-2*.

### Human pluripotent stem cells differentiation into pancreatic endocrine cells mimics human fetal endocrine cell development

To determine if similar cell populations were present in pancreatic endocrine cells derived from three hPSCs (SA121 hESC, AD2.1 iPSC, AD3.1 iPSC), these cells were differentiated to the corresponding stage (stage 5) where significant endocrine induction occurs (**Figure 7A**). CD142 and ECAD characterized three distinct populations: ECAD<sup>+</sup>CD142<sup>+</sup>, ECAD<sup>+</sup>CD142<sup>-</sup> and ECAD<sup>low</sup>CD142<sup>-</sup> as it is the case in the human fetal pancreas (9.4WD) (**Figure 7B,C**). Similar to the human fetal pancreas, SUSD2<sup>+</sup> cells were enriched in the ECAD<sup>low</sup>CD142<sup>-</sup> population (**Figure 7B,C**). Moreover, *NEUROG3*, *NEUROD1* and *NKX2-2* were mainly expressed in the ECAD<sup>low</sup>CD142<sup>SUSD2</sup><sup>+</sup> population as observed in the human fetal pancreas at 8.6WD (**Figures 6A,C** and **7D**). Finally, from stage 2–5 the temporal expression pattern of CD142 and ECAD was reminiscent of the ones occurring in the human fetal pancreas with a sharp decrease of CD142<sup>+</sup>SUSD2<sup>-</sup> subsets (from  $95 \pm 2$  of CD142<sup>+</sup>SUSD2<sup>-</sup> at stage 2 to  $23 \pm 4$  of CD142<sup>+</sup>SUSD2<sup>-</sup> subset;  $p < 0.05$ ) (**5D** and **7E**), as was also the case for SUSD2 expression (**Figures 5E** and **7F**) in the three hPSC lines. To conclude we demonstrate that pancreatic endocrine cells derived from human pluripotent stem cells appear to go through the same intermediate developmental stages as observed during in vivo development.

## Discussion

In this study, we reconstruct human fetal pancreatic differentiation by combining a specific combination of cell surface markers. Although a large set of data is available concerning development in rodent models, there is only limited knowledge on human development. Importantly, while rodent and human pancreatic development share many similarities, they differ on several aspects (*Scharfmann et al., 2013; Nair and Hebrok, 2015*). For example, the global shape and the way islet cells cluster are different between rodent and human pancreas (*Brissova et al., 2005*). Moreover, the expression pattern of major transcription factors crucial for proper pancreas development such



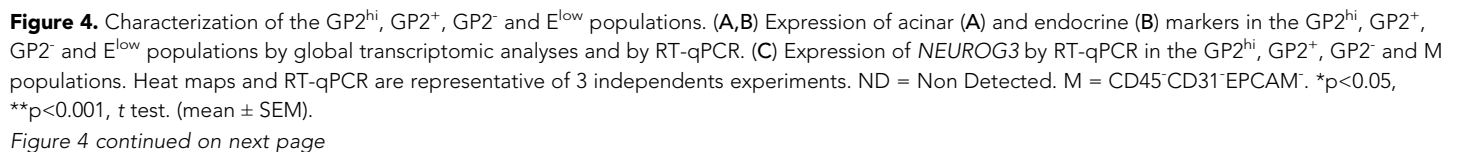




Figure 4 continued

DOI: [10.7554/eLife.27564.012](https://doi.org/10.7554/eLife.27564.012)

The following figure supplement is available for figure 4:

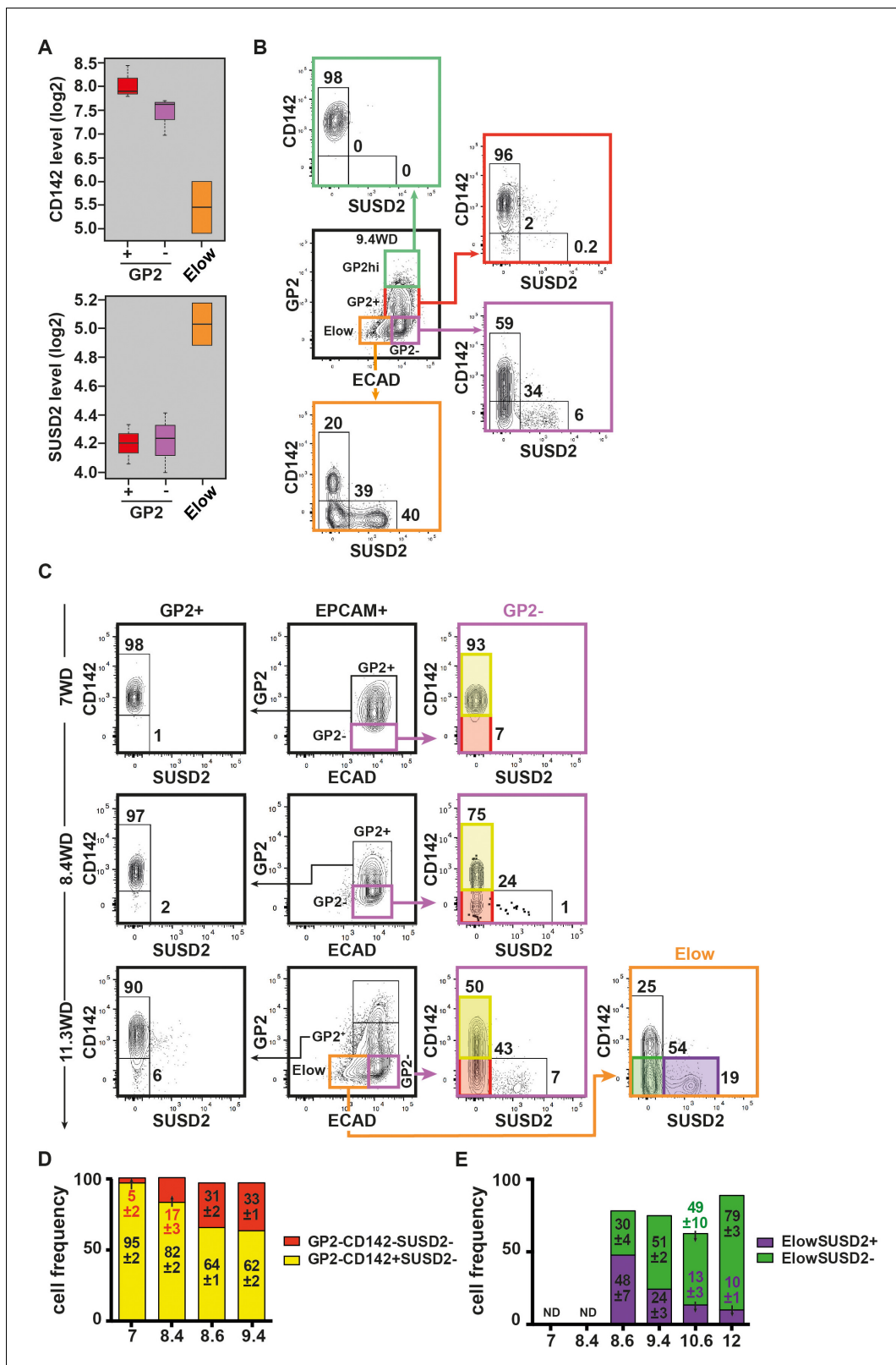
**Figure supplement 1.** Expression of ductal markers in the GP2<sup>hi</sup>, GP2<sup>+</sup>, GP2<sup>-</sup> and E<sup>low</sup> populations.

DOI: [10.7554/eLife.27564.013](https://doi.org/10.7554/eLife.27564.013)

as NEUROG3, NKX2-2 and PDX1 differ between rodent and human (Villasenor et al., 2008; Salisbury et al., 2014; Jennings et al., 2013; Sussel et al., 1998; Ohlsson et al., 1993; Heimberg et al., 2000). As an example, NEUROG3 is expressed in two waves during rodent fetal pancreatic development (Villasenor et al., 2008) while only a single-phase is observed in human (Jennings et al., 2013). It can be speculated that the number of differences between rodents and humans are under-estimated and more will be discovered by studying human pancreatic development. This knowledge will be essential to better design protocols to direct pluripotent stem cells towards functional insulin-producing pancreatic beta cells (Pagliuca et al., 2014; Rezania et al., 2014; Russ et al., 2015).

Major progress has recently been made in the field of hPSC differentiation and a number of different cell types can be efficiently generated either in vivo following transplantation into immune-incompetent mice or in vitro under controlled conditions. However, the cells generated in vitro from hPSC do not seem fully differentiated and in some cases, display a fetal phenotype rather than an adult one. This is the case for hepatocytes (Baxter et al., 2015), cardiomyocytes (Karakikes et al., 2015), neurons (Playne and Connor, 2017) and pancreatic beta cells (Hrvatin et al., 2014). These partial successes could be in part due to the limited knowledge on human cell development. Here, we specifically designed new approaches to dissect in great detail pancreatic differentiation in human. We based our study on human fetal pancreases from 7 to 13WD. This developmental period corresponds to E12.5-E17 in the mouse when pancreatic progenitors first proliferate and next develop into differentiated cells (Gittes, 2009; Jennings et al., 2015). At 6-7WD, the human pancreatic epithelium is mainly composed of PDX1<sup>+</sup>NKX6-1<sup>+</sup> pancreatic progenitors (Cebola et al., 2015; Riedel et al., 2012) while endocrine cells are extremely rare (Polak et al., 2000; Castaing et al., 2001). During the following weeks, proliferating epithelial pancreatic progenitors differentiate into endocrine, acinar and duct cells (Jennings et al., 2013; Capito et al., 2013).

The fetal pancreas is a compound organ composed of epithelial, mesenchymal, endothelial and hematopoietic cells. We excluded endothelial and hematopoietic cells using CD31 and CD45 antibodies and used the transmembrane glycoprotein EPCAM as a marker to segregate the fetal pancreatic epithelium from the mesenchyme. Previous data indicated that EPCAM is expressed during human fetal life (18-20WD) in the ductal pancreatic epithelium and in developing islet-like cells but also in the adult human pancreas in duct and islet cells (Cirulli et al., 1998). Based on this previous report we demonstrated that EPCAM also marks the fetal pancreatic epithelium at earlier developmental stages, between 7 and 12WD. Moreover, we demonstrated that the EPCAM<sup>+</sup> compartment contains PDX1<sup>+</sup>NKX6-1<sup>+</sup> double-positive cells, a hallmark of pancreatic progenitors (Jennings et al., 2013; Cebola et al., 2015). Our data also indicate that 45% of the EPCAM<sup>+</sup> compartment expresses lower levels of PDX1 and NKX6-1. It would be interesting to determine if these cells are upstream progenitors of EPCAM<sup>+</sup>PDX1<sup>+</sup>NKX6-1<sup>+</sup> cells as recently suggested (Ameri et al., 2017). Then we further segregated the fetal pancreatic epithelium into four distinct populations using the cell surface markers GP2 and ECAD. GP2 is a glycoprotein that is highly enriched in the acinar cells of the adult pancreas (Hoops and Rindler, 1991; Yu et al., 2004). Limited knowledge is available on GP2 expression during pancreatic development. Very recently, using a model of pluripotent stem cells differentiation into pancreatic endocrine cells, GP2 was identified as a novel cell surface marker of human pancreatic progenitors (Ameri et al., 2017). Our ex vivo data further support this claim. First, we observed that at 7WD, nearly all pancreatic epithelial cells are GP2<sup>+</sup>. Moreover, as development proceeds, the frequency of GP2<sup>+</sup> cells decreases. Finally, transcriptomic profiling performed at different developmental stages strongly suggests that the GP2<sup>+</sup> population can differentiate into endocrine and exocrine cells. Collectively, these data indicate that in the early human fetal pancreas, GP2 is indeed a cell surface marker of a multipotent cell population. Although the GP2<sup>+</sup> cell population is multipotent, it is not yet certain if the GP2<sup>+</sup> population contains multipotent GP2<sup>+</sup> progenitor that



**Figure 5.** Expression of CD142 and SUSD2 in the GP2<sup>hi</sup>, GP2<sup>+</sup>, GP2<sup>-</sup> and E<sup>low</sup> populations. (A) Expression of *CD142* and *SUSD2* in GP2<sup>+</sup>, GP2<sup>-</sup> and E<sup>low</sup> populations at 9WD by microarray analysis. Boxplots were obtained using standardized log2 expression values. (B) Expression of CD142 and SUSD2 in the GP2<sup>hi</sup> (in green), GP2<sup>+</sup> (in red), GP2<sup>-</sup> (in purple) and E<sup>low</sup> (in orange) populations at 9.4WD by flow cytometry. GP2 and ECAD expressions were gated on live CD45<sup>+</sup>CD31<sup>+</sup>EPCAM<sup>+</sup> cells. (C) Expression of CD142 and SUSD2 in the GP2<sup>+</sup> (left plot), GP2<sup>-</sup> (right plot) and E<sup>low</sup> (far right plot) populations at 7WD, 8.4WD and 11.3WD by flow cytometry. (D) Cell frequency of GP2-CD142-SUSD2- (red) and GP2-CD142+SUSD2- (yellow) populations at 7, 8.4, 8.6 and 9.4 WD. (E) Cell frequency of E<sup>low</sup>SUSD2<sup>+</sup> (purple) and E<sup>low</sup>SUSD2<sup>-</sup> (green) populations at 7, 8.4, 8.6, 9.4, 10.6 and 12 WD. ND: Not Determined.

Figure 5 continued on next page

Figure 5 continued

populations at 7WD, 8.4WD, and 11.3WD. (D) Cell frequencies of the GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>-</sup> (named CD142<sup>+</sup>SUSD2<sup>-</sup> in yellow) and the GP2<sup>+</sup>CD142<sup>-</sup>SUSD2<sup>-</sup> (CD142<sup>-</sup>SUSD2<sup>-</sup> in red) subsets from 7WD to 9.4WD. (mean  $\pm$  SEM). (E) Cell frequencies of the E<sup>low</sup>GP2<sup>+</sup>CD142<sup>-</sup>SUSD2<sup>-</sup> (named E<sup>low</sup>SUSD2<sup>-</sup>, in purple) and the E<sup>low</sup>GP2<sup>-</sup>CD142<sup>-</sup>SUSD2<sup>-</sup> (E<sup>low</sup>SUSD2<sup>-</sup>, in green) subsets from 7 to 12WD. (mean  $\pm$  SEM). (A) n = 2–3, (B) n = 3; (C) 7WD n = 2, 8.4WD to 9.4WD n = 3. (D) n = 3.

DOI: [10.7554/eLife.27564.014](https://doi.org/10.7554/eLife.27564.014)

The following figure supplement is available for figure 5:

**Figure supplement 1.** Expression of CD142 and SUSD2 in the GP2<sup>+</sup>, GP2<sup>-</sup> and E<sup>low</sup> populations at 8.6WD.

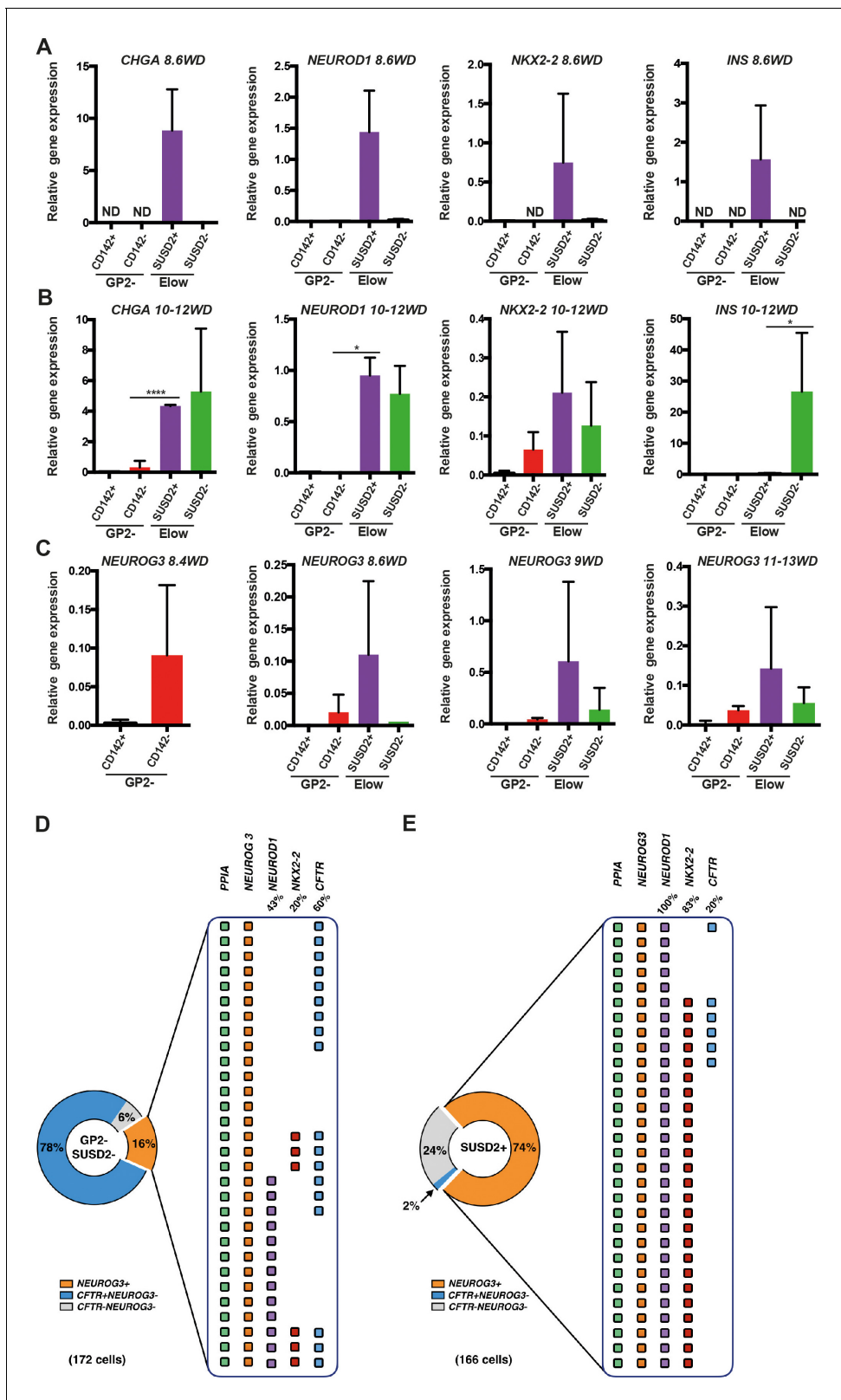
DOI: [10.7554/eLife.27564.015](https://doi.org/10.7554/eLife.27564.015)

can differentiate into the three lineages (endocrine, acinar and ductal) or represent a mixture of GP2<sup>+</sup> progenitors already committed to a specific lineage. This could be addressed by utilizing a single cell culture approach. However, our attempts so far to culture so few cells sorted from human fetal pancreas have failed. Discovering new and efficient culture conditions such as co-culture on feeder layers could alleviate this issue (Trott et al., 2017).

In animal models, cell adhesion molecules have emerged as key regulators of embryonic morphogenesis and this topic has been extensively studied using pancreatic development as a model organ (Semb, 2004). As an example, during mouse and chick development, Pdx1<sup>+</sup>Nkx6.1<sup>+</sup> progenitors express high levels of ECAD. NEUROG3<sup>+</sup> endocrine progenitors will develop from such pancreatic progenitors while lowering their ECAD level during their delamination from the ductal tree to develop into pancreatic endocrine cells (Gouzi et al., 2011). Information remains scarcer on the regulation of ECAD levels during human development. Our results indicate that at 7WD, GP2<sup>+</sup> pancreatic progenitors express ECAD at high levels. At 8.6WD, ECAD<sup>low</sup> cells appear, their frequency increasing while development progresses. Interestingly, NEUROG3-expressing cells first appear at 8.4WD in the ECAD<sup>+</sup> population and are found later on in the ECAD<sup>low</sup> population. Whether it is linked to their delamination remains to be demonstrated. This step is followed (at 8.6WD) by the expression of endocrine markers such as CHGA, NEUROD1 and INS. Of importance, endocrine cells expressed low but significant levels of ECAD. This perfectly fits with mouse data that indicate that during development ECAD function is necessary for proper aggregation of endocrine cells after delamination (Dahl et al., 1996). Taken together, our data demonstrate that ECAD levels are tightly regulated during specific steps of human pancreatic development.

We further refined specific pancreatic cell populations by using CD142 and SUSD2 found in our transcriptomic analysis as additional cell surface markers. CD142 has been proposed as a marker of pancreatic endodermal cells that also labels additional cell types (Kelly et al., 2011). SUSD2 was previously used as a marker to enrich NEUROG3<sup>+</sup> from hPSC derived pancreatic cells and the human fetal pancreas (Liu et al., 2014). Here, with our set of markers we reconstructed human pancreatic cell differentiation (Figure 7G). The full combination of markers was required for this reconstruction as none of the markers was specific to the different subsets (Figure 7—figure supplement 1). Our data indicate that GP2<sup>+</sup>CD142<sup>+</sup> pancreatic progenitors can either give rise to GP2<sup>hi</sup>CD142<sup>+</sup> acinar cells or enter the endocrine pathway and express NEUROG3 by turning off GP2 and CD142. Endocrine maturation further progresses by up-regulating SUSD2 and decreasing ECAD level (Figure 7G). The first INS<sup>+</sup> cells were detected at 8.6WD in the E<sup>low</sup> population as SUSD2<sup>+</sup> and later on as SUSD2<sup>-</sup>. Recently, differences in gene expression and functionality were observed between fetal, neonatal and adult beta cells (Hrvatin et al., 2014; Jermendy et al., 2011; Blum et al., 2012). Moreover, heterogeneity between adult beta cells was also recently described (Dorrell et al., 2016; Bader et al., 2016). Whether INS<sup>+</sup> cells positive for SUSD2 represent a first wave of beta cells that could be poly-hormonal remains to be tested. Moreover, comparative analyses between INS<sup>+</sup> cells from E<sup>low</sup>SUSD2<sup>+</sup> and E<sup>low</sup>SUSD2<sup>-</sup> subsets will define the differences in their transcriptional factor networks.

FACS-based approaches have been used since the eighties in the hematopoietic field to dissect hematopoiesis (Spangrude et al., 1988). However, it has rarely been used in the field of pancreatic development and even less frequently with human fetal pancreases. In the pancreas, the majority of the cell sorting approaches was performed by analyzing fluorescent signals from tagged proteins derived from transgenic mice (Gu et al., 2004; Miyatsuka et al., 2014). While highly informative, this strategy cannot be used to dissect human pancreatic development. Antibodies against cell



**Figure 6.** Molecular characterization of the GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>+</sup> and E<sup>low</sup>GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>+</sup> subsets. (A, B) RT-qPCR for *CHGA*, *NEUROD1*, *NKX2-2* and *INS* on sorted GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>+</sup> (GP2<sup>+</sup>CD142<sup>+</sup>), GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>-</sup> (GP2<sup>+</sup>CD142<sup>-</sup>), E<sup>low</sup>GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>+</sup> (E<sup>low</sup>SUSD2<sup>+</sup>) and E<sup>low</sup>GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>-</sup> (E<sup>low</sup>SUSD2<sup>-</sup>) subsets at 8.6 and 10-12WD. (C) RT-qPCR for *NEUROG3* at 8.4, 8.6, 9 and 11-13WD in the GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>+</sup> (GP2<sup>+</sup>CD142<sup>+</sup>), GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>-</sup> (GP2<sup>+</sup>CD142<sup>-</sup>), E<sup>low</sup>GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>+</sup> (E<sup>low</sup>SUSD2<sup>+</sup>) and E<sup>low</sup>GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>-</sup> (E<sup>low</sup>SUSD2<sup>-</sup>) subsets. (D, E) Single cell RT-qPCR

Figure 6 continued on next page

Figure 6 continued

at 9WD on 172 GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>-</sup> (GP2<sup>+</sup>SUSD2<sup>-</sup>, left panel) and 166 E<sup>low</sup>GP2<sup>+</sup>CD142<sup>+</sup>SUSD2<sup>+</sup> (SUSD2<sup>+</sup>, right panel) cells for the expression of *PPIA*, *NEUROG3*, *NEUROD1*, *NKX2-2* and *CFTR*. Pie charts represent the percentage of *NEUROG3*<sup>+</sup> (in orange), *NEUROG3*<sup>+</sup>*CFTR*<sup>+</sup> (in blue) and *NEUROG3*<sup>+</sup>*CFTR*<sup>-</sup> cells (in grey). For *NEUROG3*<sup>+</sup> the percentages of *NEUROD1*<sup>+</sup>, *NKX2-2*<sup>+</sup>, and *CFTR*<sup>+</sup> cells are displayed. Each line represents one cell. *PPIA*<sup>-</sup> cells were excluded from the analysis. (A–C) n = 3, (D, E) n = 2. \*p<0.05, \*\*\*\*p<0.0001, t test. (mean ±SEM).

DOI: [10.7554/eLife.27564.016](https://doi.org/10.7554/eLife.27564.016)

The following figure supplement is available for figure 6:

**Figure supplement 1.** Expression of *CHGA* and *NEUROG3* in the E<sup>low</sup>CD142<sup>+</sup> subsets.

DOI: [10.7554/eLife.27564.017](https://doi.org/10.7554/eLife.27564.017)

surface markers were used in a limited number of studies on pancreas development and mainly in rodents. It was found that a combination of CD49f and CD133 antibodies can be used to enrich fraction in mouse pancreatic progenitors expressing *NEUROG3*-expressing cells with some data on human fetal pancreas. However, the enrichment in *NEUROG3* on human fetal pancreas was limited using this combination of markers (Sugiyama et al., 2007). More recently, SUSD2 was used as a marker to enrich *NEUROG3*<sup>+</sup> cells in hPSC -derived pancreatic cells and in the human fetal pancreas (Liu et al., 2014). Our data confirm this point. However our single cell qPCR indicate that SUSD2 does not mark all, but only a subset of *NEUROG3*<sup>+</sup> that co-expressed *NEUROD1* and *NKX2-2*. Thus, by using a combination of antibodies against cell surface markers, we demonstrate that cell populations highly enriched in specific functions can be sorted from the human fetal pancreas at different stages of development, allowing the reconstruction of the differentiation program. Our model constitutes a key advancement in understanding human fetal pancreas development by mapping out the pattern of differentiation of the three main pancreatic lineages. We further refined our work for the endocrine pathway by describing for the first time, discrete stages of human pancreatic endocrine cell differentiation and showed that our development model also applies to the in vitro differentiation of hPSCs into pancreatic endocrine cells. To the best of our knowledge, this type of side-by-side comparison that demonstrates that in vitro hPSC differentiation mimics in vivo events has rarely been done.

In conclusion, we provide a novel way of approaching human pancreatic differentiation. Our work will be useful to fill the limited knowledge on human pancreas development. It should also pave the way for developing new cell therapies for diabetic patients.

## Materials and methods

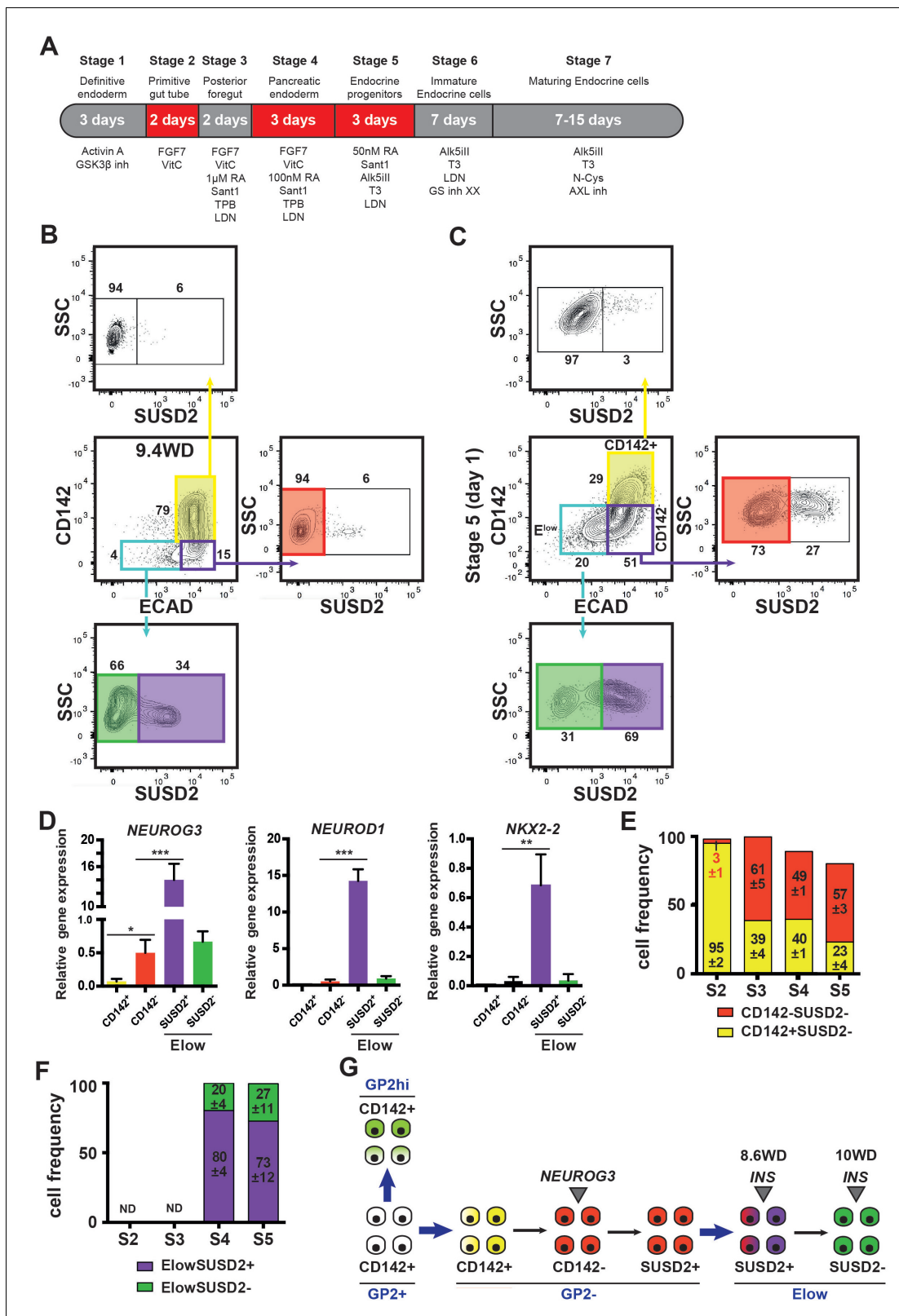
### Pancreatic dissection and cell suspension preparation

All experiments on human fetal pancreas were performed at INSERM Paris, France. Human fetal pancreases were isolated from surgical abortion done by suction aspiration between 7 to 13 weeks of development (Castaing et al., 2001; Capito et al., 2013; Scharfmann et al., 2014) in compliance with the French bioethics legislation and the guidelines of our institution. Approval was obtained from Agence de Biomedecine, the French competent authority along with maternal written consent. Ages were determined on the basis of time since the last menstrual period and hand and foot morphology. Fetal pancreases were micro dissected with forceps under a binocular magnifying lens, rinsed with Hanks Balanced Salt Solution (HBSS) from Gibco to remove contaminating blood cells and gently disrupted using forceps. Afterwards, pancreases were incubated for 5 min in collagenase V (0.5 mg/ml) (Sigma Aldrich) in HBSS in the presence of calcium and magnesium. Cells were rinsed in HBSS and then incubated for 5 min in trypsin (0.05%) (Gibco). Finally, cells were rinsed in HBSS supplemented with 20% Fetal calf serum (from Eurobio). Pluripotent stem cells were washed with PBS without Ca<sup>2+</sup> and Mg<sup>2+</sup> (Invitrogen, FRANCE) and incubated with TrypLE select for 1–3 min.

### Maintenance and differentiation of human pluripotent stem cell lines

A human ESC (SA121) (Heins et al., 2004) obtained from Takara and two iPSC (SB Ad2.1 and SB3.1) (van de Bunt et al., 2016) obtained from the StemBANCC consortium were applied for this study. These cell lines had been confirmed to be pluripotent by evaluation of pluripotency marker expression, tri-lineage differentiation and karyotyping and tested negative for mycoplasma contamination.





**Figure 7.** Expression of CD142, ECAD and SUSD2 in pancreatic endocrine cells derived from hPSCs. (A) Scheme of in vitro differentiation of hPSCs into endocrine pancreatic cells. (B, C) Expression of CD142, ECAD (gated on lived CD45<sup>+</sup>CD31<sup>+</sup>EPCAM<sup>+</sup> cells for the human) and SUSD2 in CD142<sup>+</sup>, CD142<sup>-</sup> and E<sup>low</sup> population by flow cytometry in the human fetal pancreas at 9.4WD (in B) and in pancreatic endocrine cells derived from hPSCs (AD3.1 iPSC) at stage 5 (in C). (D) RT-qPCR on CD142<sup>+</sup>E-CAD<sup>+</sup>SUSD2<sup>-</sup> (named CD142<sup>+</sup>), CD142<sup>+</sup>E-CAD<sup>+</sup>SUSD2<sup>+</sup> (named CD142<sup>+</sup>), CD142<sup>+</sup>E-CAD<sup>low</sup>SUSD2<sup>+</sup> (named CD142<sup>low</sup>) and CD142<sup>+</sup>E-CAD<sup>low</sup>SUSD2<sup>-</sup> (named CD142<sup>low</sup>) in human fetal pancreas and hPSC-derived cells. (E) Cell frequency of CD142-SUSD2<sup>-</sup> (red) and CD142+SUSD2<sup>-</sup> (yellow) in hPSC-derived cells. (F) Cell frequency of ElowSUSD2<sup>+</sup> (purple) and ElowSUSD2<sup>-</sup> (green) in hPSC-derived cells. (G) Schematic of cell fate transitions. GP2<sup>hi</sup> CD142<sup>+</sup> cells transition to GP2<sup>-</sup> CD142<sup>+</sup> cells, then to GP2<sup>-</sup> CD142<sup>-</sup> cells, and finally to Elow SUSD2<sup>+</sup> and Elow SUSD2<sup>-</sup> cells. Key factors include NEUROG3, INS, and GP2<sup>+</sup>.



Figure 7 continued

E<sup>low</sup>SUSD2<sup>+</sup>), CD142<sup>+</sup>E-CAD<sup>low</sup>SUSD2<sup>+</sup> (named E<sup>low</sup>SUSD2<sup>+</sup>) from hPSCs (AD3.1 iPSC) at stage five for *NEUROG3*, *NEUROD1* and *NKX2-2*. (E) Cell frequencies of CD142<sup>+</sup>SUSD2<sup>+</sup> (in yellow) and CD142<sup>+</sup>SUSD2<sup>+</sup> (in red) from hPSCs (SA121 hESC, AD2.1 iPSC, AD3.1 iPSC) from stage 2–5. (F) Cell frequencies of E<sup>low</sup>SUSD2<sup>+</sup> and E<sup>low</sup>SUSD2<sup>+</sup> from hPSCs (SA121 hESC, AD2.1 iPSC, AD3.1 iPSC) from stage 2–5. (G) Scheme representing human pancreatic differentiation across development using cell surface markers. (B–F) n = 3. \*p<0.05, \*\*p<0.001, \*\*\*p<0.005, t test.

DOI: 10.7554/eLife.27564.018

The following figure supplement is available for figure 7:

**Figure supplement 1.** Expression of CD142 and SUSD2 in the mesenchyme, the endothelial/hematopoietic and the epithelial compartments.

DOI: 10.7554/eLife.27564.019

All three lines were cultured in mTeSR1 medium (StemCell Technologies) on hESC-qualified matrigel (Corning). Cells were passaged every 3–4 days or when confluent by dissociating to a single cell solution using TrypLE select (ThermoFisher). Single cells were seeded onto freshly coated Matrigel tissue culture flask in mTeSR1 containing of 5  $\mu$ M Tiger (Rock inhibitor, Sigma-Aldrich) and medium was replenished daily. For differentiation of the hESC/hiPSC lines cells were dissociated to a single cell solution using TrypLE select and resuspended mTeSR1 with 5  $\mu$ M Tiger. Cells were seeded at a concentration of  $0.35 \times 10^6$  cells/cm<sup>2</sup> onto growth-factor reduced matrigel CellBIND surfaces (Corning). Cells were incubated for 24 hr before the start of the differentiation. Differentiation to the pancreatic lineage was conducted as described in a previously published protocol (Rezania et al., 2014) with the following modifications: CHIR99201 (Axon Medchem) was applied at 3  $\mu$ M and 0.3  $\mu$ M concentration for the first and second day, respectively of the definitive endoderm differentiation instead of MCX-928 used in the original protocol. MCDB131 medium (Life technologies) was used as basal medium throughout the differentiation instead of BLAR medium. Cells were not dissociated and incubated as clusters on air-liquid filters during stage 5; instead cells were kept in 2D cultures throughout the differentiation.

## Flow cytometry

Following dissociation, cells were incubated with antibodies for 20 min in FACS medium (HBSS +2% FCS), then rinsed in FACS medium and re-suspended in FACS medium with Propidium Iodide (1/4000) (Sigma Aldrich) or DAPI solution (0,1  $\mu$ g/ml) (BD Biosciences) to label dead cells. For intra-cellular staining cells were fixed for 5 min in 3% PFA and then rinsed in DPBS from GIBCO. Then cells were permeabilized in DPBS +3% BSA +0.3% Triton and incubated with antibodies overnight at 4°C, rinsed and re-suspended in DPBS. For each antibody, optimal dilution was determined by titration. The following antibodies were used: anti CD45-PerCP/Cy5.5 (1/20, clone 2D1, Biolegend; RRID: AB\_2566351), anti CD31-PerCP/Cy5.5 (1/20, clone WM59, Biolegend; RRID: AB\_2566174), anti-EPCAM-Brilliant violet 605 (1/20, clone 9C4, Biolegend; RRID: AB\_2562518), anti ECAD-PE-Cy7 (1/20, clone 67A4, Biolegend; RRID: AB\_2563096), anti-GP2-PE (1/5, clone 3G7H9, MBL; RRID: AB\_11160953), anti-SUSD2-VioBrightFITC (1/20, W5C5, Miltenyi Biotec; RRID: AB\_2653618), anti-CD142-VioBlue (1/20, clone HTF-1, Miltenyi Biotec; RRID: AB\_2655132), anti-PDX1-alexaFluor 488 (1/20, clone 658A5, BD Biosciences) and anti-NKX6-1-alexaFluor 647 (1/20, clone R11-560, BD Biosciences). An ARIA III (BD Bioscience) was used for cell sorting and a FACS LSRFortessa for analysis (BD Bioscience). Data were analyzed using FlowJo 10.2 software. Dead cells were excluded from analyses performed on lived cells. Lived cells referred to cells that did not incorporate the dead cell stain Propidium Iodide.

## Immunohistochemistry

Human fetal pancreatic sections (4–5  $\mu$ m thick) were prepared and processed as previously described (Castaing et al., 2005). The following primary antibodies were used anti-mouse EPCAM (1:500, clone VU1D9, Cell Signaling) and anti-rabbit PDX1 (1/1000) (Duvill   et al., 2003). The secondary antibodies were anti-mouse Alexa Fluor 488 antibodies (1:400, Life Technologies) and anti-rabbit Alexa Fluor 594 antibodies (1:400, Jackson ImmunoResearch). The nuclei were stained using the Hoechst 33342 fluorescent stain (0.3 mg/ml, Invitrogen, France).

## Bulk and single cells RT-qPCR

For human fetal pancreas, single to 100 cells were sorted in 9  $\mu$ L of RT/pre-amp mix from the One-Step qRT-PCR Kit (Invitrogen, France). Pre-amplified (20 cycles) cDNA was obtained according to manufacturer's notice and was used for qPCR reaction. The pre-amplified cDNA was then used for qPCR using the TaqMan protocol. RT, pre-amplification and qPCR were performed using TaqMan primers from Applied Biosystems. The following TaqMan primers were used: *PPIA* (Hs04194521\_s1), *HPRT1* (Hs99999909\_m1), *PDX1* (Hs00236830\_m1 and Hs00426216\_m1), *NKX6-1* (Hs00232355\_m1), *CEL* (Hs01068709\_m1), *CELA2a* (Hs04194660\_s1), *CELA3A* (Hs00371667\_gH), *CTRC* (Hs00200713\_m1), *SOX9* (Hs00165814\_m1), *NEUROG3* (Hs01875204\_s1), *CHGA* (Hs00900370\_m1), *NEUROD1* (Hs01922995\_s1), *INS* (Hs02741908), *NKX2-2* (Hs00159616\_m1). RT-qPCR results are presented in arbitrary units (AU) relative to expression of the control gene *PPIA* (for human qPCR) and *HPRT* (for hPSC qPCRs). All qPCRs on the human fetal pancreas were assayed using *PPIA* while *HPRT* was used for the qPCR on hPSCs. QPCRs were run on QuantStudio from ThermoFischer and single cell qPCR were assayed on the Biomark from Fluidigm following manufacture's instructions.

For human pluripotent stem cell differentiation cells were sorted from three independent differentiations of the ADSB3.1 iPSC line. Cells were lysed in RP1 lysis buffer from the NucleoSpin RNA/Protein purification kit (Macherey-Nagel) and stored at  $-80^{\circ}\text{C}$  until purification of RNA. RNAs were extracted using NucleoSpin RNA/protein purification kit. RNA was subsequently converted to cDNA using iScript cDNA synthesis kit (Biorad) according to manufactures instructions. RT-qPCR was performed on a MX3005P qPCR system (Agilent Genomics) using a fast-2-step protocol (first  $95^{\circ}\text{C}$  for 1 min, then 40 cycles of  $95^{\circ}\text{C}$  for 10 s,  $60^{\circ}\text{C}$  for 25 s). The following TagMan primers were used: *NEUROG3* (Hs01875204\_s1), *NEUROD1* (Hs01922995\_s1), *NKX2-2* (Hs00159616\_m1), *bActin* (Hs01060665\_g1) and *HPRT* (Hs99999909\_m1).

## Global transcriptomic analysis

Total RNA was obtained from Trizol-preserved (Invitrogen, France) sorted cells by chloroform extraction, quantified and assessed for quality by Agilent-2100 Bioanalyzer (Agilent, Santa Clara, CA). cDNAs were prepared and amplified using Ovation Pico WTA System V2 kit (NuGEN Technologies, San Carlos, CA, USA) and hybridized onto GeneChip Human Gene 2.0 ST Array (Affymetrix, Santa Clara, CA, USA). Quality Control and normalization (RMA) were conducted using Bioconductor Project software (<https://www.bioconductor.org>), which provides log<sub>2</sub> transformed expression values. Data were extracted from raw CEL files using a custom GeneChip library file (CDF file) provided by <http://brainarray.mbni.med.umich.edu/CustomCDF> (Dai et al., 2005).

## Transcriptomic statistical analysis

Two-sample comparisons were made using two-tailed Student t-test. Transcripts for which any p-value was above 0.05 were filtered out, leaving 6444 transcripts for further analyses. We compared expression of these genes in each cell population using Gene Set Enrichment Analysis (GSEA) software and the 'Biological Process' database of the Gene Ontology Consortium (Subramanian et al., 2005). Biological processes with FDR q-value <0.05 were considered significant.

Cut off to define the 'specific enriched genes':

Each population ( $\text{GP2}^{\text{hi}}$ ,  $\text{GP2}^+$ ,  $\text{GP2}^-$ ,  $\text{E}^{\text{low}}$  and M) was compared to the other populations using two-tailed Student t-test. Genes were considered as enriched in a specific population when over-expressed (p-value<0,05 and Fold Change > 2) in a population compared to any of the others. The list of the 'specific enriched genes' is displayed in **supplementary file 1b and c**.

Expression pattern of the enriched fetal genes in the adult pancreas:

We use data from the single-cell RNA-seq (Segerstolpe et al., 2016). The normalized data are available on ArrayExpress (<http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-5060>). Next, we generated heatmaps displaying the expression of the fetal 'specific enriched genes' in the adult acinar, alpha, beta, epsilon, gamma and ductal cells.

Heatmaps were generated by the 'heatmap2' function from gplots R package (<https://cran.r-project.org/web/packages/gplots/>) on standardized log<sub>2</sub> expression values, with Pearson correlation as the distance function.

## Acknowledgements

We would like to thank Delphine Bredele (INSERM U1016) for technical assistance and Ana Cumano (INSERM U668) for supports. This work was supported by the HumEn project funded by the European Commission's Seventh Framework Programme for Research, (agreement No 602587) (RS and HS) and the Foundation Bettencourt Schueller (to RS). The RS laboratory belongs to the Laboratoire d'Excellence consortium Revive and to the Departement Hospitalo-Universitaire (DHU) Autoimmune and Hormonal disease. The research leading to the results on the hiPSC has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement n° 115439, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. This publication reflects only the authors' views and neither the IMI JU nor EFPIA nor the European Commission are liable for any use that may be made of the information contained therein. Work on the hESC line was done without the support from StemBANCC.

## Additional information

### Funding

Funder	Grant reference number	Author
Seventh Framework Programme	602587	Raphaël Scharfmann
Innovative Medicines Initiative	115439	Raphaël Scharfmann

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

CR, Conceptualization, Formal analysis, Investigation, Methodology, Writing—original draft; NG, Formal analysis, Methodology; CB, Investigation, Methodology; JA, Conceptualization, Writing—review and editing; JSK, Investigation, Writing—review and editing; MH, Supervision, Funding acquisition, Writing—review and editing; CH, Investigation, Methodology, Writing—review and editing; HS, Conceptualization, Supervision, Writing—review and editing; RS, Conceptualization, Funding acquisition, Writing—original draft

### Author ORCIDs

Raphaël Scharfmann,  <http://orcid.org/0000-0001-7619-337X>

## Additional files

### Supplementary files

- Supplementary file 1. Gene Ontology list. (a) List of gene sets and corresponding genes used in GSEA analysis displayed in **Figure 3B**. Gene sets were obtained from the gene ontology database. (b) List of differentially expressed genes specific to each population at 9WD. E<sup>low</sup>: in yellow, GP2<sup>-</sup>: in purple, GP2<sup>+</sup>: in red, M: white. Values are displayed in Log2. (c) List of differentially expressed genes specific to each population at 11WD. E<sup>low</sup>: in yellow, GP2<sup>-</sup>: in purple, GP2<sup>+</sup>: in red, M: white. Values are displayed in Log2.

DOI: [10.7554/eLife.27564.020](https://doi.org/10.7554/eLife.27564.020)

### Major datasets

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database, license, and accessibility information
Ramond C, Glaser N, Berthault C,	2017	Expression data from cells sorted from human fetal pancreas	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc">https://www.ncbi.nlm.nih.gov/geo/query/acc</a> .	Publicly available at NCBI Gene Expression

Ameri J, Kirkegaard JS, Hansson M, Honoré C, Semb H, Scharfmann R

cgi?acc=GSE96697

Omnibus (accession no: GSE96697)

## References

- Ameri J, Borup R, Prawiro C, Ramond C, Schachter KA, Scharfmann R, Semb H. 2017. Efficient generation of glucose-responsive beta cells from isolated GP2(+) Human pancreatic progenitors. *Cell Reports* **19**:36–49. doi: [10.1016/j.celrep.2017.03.032](https://doi.org/10.1016/j.celrep.2017.03.032), PMID: [28380361](https://pubmed.ncbi.nlm.nih.gov/28380361/)
- Bader E, Migliorini A, Gegg M, Moruzzi N, Gerdes J, Roscioni SS, Bakhti M, Brandl E, Irmeler M, Beckers J, Aichler M, Feuchtinger A, Leitzinger C, Zischka H, Wang-Sattler R, Jastroch M, Tschöp M, Machicao F, Staiger H, Häring HU, et al. 2016. Identification of proliferative and mature  $\beta$ -cells in the islets of langerhans. *Nature* **535**:430–434. doi: [10.1038/nature18624](https://doi.org/10.1038/nature18624), PMID: [27398620](https://pubmed.ncbi.nlm.nih.gov/27398620/)
- Baxter M, Withey S, Harrison S, Segeritz CP, Zhang F, Atkinson-Dell R, Rowe C, Gerrard DT, Sison-Young R, Jenkins R, Henry J, Berry AA, Mohamet L, Best M, Fenwick SW, Malik H, Kitteringham NR, Goldring CE, Piper Hanley K, Vallier L, et al. 2015. Phenotypic and functional analyses show stem cell-derived hepatocyte-like cells better mimic fetal rather than adult hepatocytes. *Journal of Hepatology* **62**:581–589. doi: [10.1016/j.jhep.2014.10.016](https://doi.org/10.1016/j.jhep.2014.10.016), PMID: [25457200](https://pubmed.ncbi.nlm.nih.gov/25457200/)
- Benthuyssen JR, Carrano AC, Sander M. 2016. Advances in  $\beta$  cell replacement and regeneration strategies for treating diabetes. *Journal of Clinical Investigation* **126**:3651–3660. doi: [10.1172/JCI87439](https://doi.org/10.1172/JCI87439), PMID: [27694741](https://pubmed.ncbi.nlm.nih.gov/27694741/)
- Bhushan A, Itoh N, Kato S, Thiery JP, Czernichow P, Bellusci S, Scharfmann R. 2001. Fgf10 is essential for maintaining the proliferative capacity of epithelial progenitor cells during early pancreatic organogenesis. *Development* **128**:5109–5117. PMID: [11748146](https://pubmed.ncbi.nlm.nih.gov/11748146/)
- Blum B, Hrvatin SS, Schuetz C, Bonal C, Rezanian A, Melton DA. 2012. Functional beta-cell maturation is marked by an increased glucose threshold and by expression of urocortin 3. *Nature Biotechnology* **30**:261–264. doi: [10.1038/nbt.2141](https://doi.org/10.1038/nbt.2141), PMID: [22371083](https://pubmed.ncbi.nlm.nih.gov/22371083/)
- Brissova M, Fowler MJ, Nicholson WE, Chu A, Hirshberg B, Harlan DM, Powers AC. 2005. Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy. *Journal of Histochemistry & Cytochemistry* **53**:1087–1097. doi: [10.1369/jhc.5C6684.2005](https://doi.org/10.1369/jhc.5C6684.2005), PMID: [15923354](https://pubmed.ncbi.nlm.nih.gov/15923354/)
- Capito C, Simon MT, Aiello V, Clark A, Aigrain Y, Ravassard P, Scharfmann R. 2013. Mouse muscle as an ectopic permissive site for human pancreatic development. *Diabetes* **62**:3479–3487. doi: [10.2337/db13-0554](https://doi.org/10.2337/db13-0554), PMID: [23835344](https://pubmed.ncbi.nlm.nih.gov/23835344/)
- Castaing M, Guerci A, Mallet J, Czernichow P, Ravassard P, Scharfmann R. 2005. Efficient restricted gene expression in beta cells by lentivirus-mediated gene transfer into pancreatic stem/progenitor cells. *Diabetologia* **48**:709–719. doi: [10.1007/s00125-005-1694-6](https://doi.org/10.1007/s00125-005-1694-6), PMID: [15759109](https://pubmed.ncbi.nlm.nih.gov/15759109/)
- Castaing M, Péault B, Basmaciogullari A, Casal I, Czernichow P, Scharfmann R. 2001. Blood glucose normalization upon transplantation of human embryonic pancreas into beta-cell-deficient SCID mice. *Diabetologia* **44**:2066–2076. doi: [10.1007/s001250100012](https://doi.org/10.1007/s001250100012), PMID: [11719839](https://pubmed.ncbi.nlm.nih.gov/11719839/)
- Cebola I, Rodríguez-Seguí SA, Cho CH, Bessa J, Rovira M, Luengo M, Chhatrivala M, Berry A, Ponsa-Cobas J, Maestro MA, Jennings RE, Pasquali L, Morán I, Castro N, Hanley NA, Gomez-Skarmeta JL, Vallier L, Ferrer J. 2015. TEAD and YAP regulate the enhancer network of human embryonic pancreatic progenitors. *Nature Cell Biology* **17**:615–626. doi: [10.1038/ncb3160](https://doi.org/10.1038/ncb3160), PMID: [25915126](https://pubmed.ncbi.nlm.nih.gov/25915126/)
- Cirulli V, Crisa L, Beattie GM, Mally MI, Lopez AD, Fannon A, Ptaszniak A, Inverardi L, Ricordi C, Deerinck T, Ellisman M, Reisfeld RA, Hayek A. 1998. KSA antigen Ep-CAM mediates cell-cell adhesion of pancreatic epithelial cells: morphoregulatory roles in pancreatic islet development. *The Journal of Cell Biology* **140**:1519–1534. doi: [10.1083/jcb.140.6.1519](https://doi.org/10.1083/jcb.140.6.1519), PMID: [9508783](https://pubmed.ncbi.nlm.nih.gov/9508783/)
- Dahl U, Sjödin A, Semb H. 1996. Cadherins regulate aggregation of pancreatic beta-cells in vivo. *Development* **122**:2895–2902. PMID: [8787762](https://pubmed.ncbi.nlm.nih.gov/8787762/)
- Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, Bunney WE, Myers RM, Speed TP, Akil H, Watson SJ, Meng F. 2005. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Research* **33**:e175. doi: [10.1093/nar/gni179](https://doi.org/10.1093/nar/gni179), PMID: [16284200](https://pubmed.ncbi.nlm.nih.gov/16284200/)
- Dorrell C, Schug J, Canaday PS, Russ HA, Tarlow BD, Grompe MT, Horton T, Hebrok M, Streeter PR, Kaestner KH, Grompe M. 2016. Human islets contain four distinct subtypes of  $\beta$  cells. *Nature Communications* **7**:11756. doi: [10.1038/ncomms11756](https://doi.org/10.1038/ncomms11756), PMID: [27399229](https://pubmed.ncbi.nlm.nih.gov/27399229/)
- Duvillié B, Attali M, Aiello V, Quemeneur E, Scharfmann R. 2003. Label-retaining cells in the rat pancreas: location and differentiation potential in vitro. *Diabetes* **52**:2035–2042. doi: [10.2337/diabetes.52.8.2035](https://doi.org/10.2337/diabetes.52.8.2035), PMID: [12882920](https://pubmed.ncbi.nlm.nih.gov/12882920/)
- Eaves CJ. 2015. Hematopoietic stem cells: concepts, definitions, and the new reality. *Blood* **125**:2605–2613. doi: [10.1182/blood-2014-12-570200](https://doi.org/10.1182/blood-2014-12-570200), PMID: [25762175](https://pubmed.ncbi.nlm.nih.gov/25762175/)
- Gittes GK. 2009. Developmental biology of the pancreas: a comprehensive review. *Developmental Biology* **326**:4–35. doi: [10.1016/j.ydbio.2008.10.024](https://doi.org/10.1016/j.ydbio.2008.10.024), PMID: [19013144](https://pubmed.ncbi.nlm.nih.gov/19013144/)
- Gouzi M, Kim YH, Katsumoto K, Johansson K, Grapin-Botton A. 2011. Neurogenin3 initiates stepwise delamination of differentiating endocrine cells during pancreas development. *Developmental Dynamics* **240**:589–604. doi: [10.1002/dvdy.22544](https://doi.org/10.1002/dvdy.22544), PMID: [21287656](https://pubmed.ncbi.nlm.nih.gov/21287656/)

- Gradwohl G**, Dierich A, LeMeur M, Guillemot F. 2000. neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *PNAS* **97**:1607–1611. doi: [10.1073/pnas.97.4.1607](https://doi.org/10.1073/pnas.97.4.1607), PMID: [10677506](https://pubmed.ncbi.nlm.nih.gov/10677506/)
- Gu G**, Dubauskaite J, Melton DA. 2002. Direct evidence for the pancreatic lineage: ngn3+ cells are islet progenitors and are distinct from duct progenitors. *Development* **129**:2447–2457. PMID: [11973276](https://pubmed.ncbi.nlm.nih.gov/11973276/)
- Gu G**, Wells JM, Dombkowski D, Pfeffer F, Aronow B, Melton DA. 2004. Global expression analysis of gene regulatory pathways during endocrine pancreatic development. *Development* **131**:165–179. doi: [10.1242/dev.00921](https://doi.org/10.1242/dev.00921), PMID: [14660441](https://pubmed.ncbi.nlm.nih.gov/14660441/)
- Heimberg H**, Bouwens L, Heremans Y, Van De Casteele M, Lefebvre V, Pipeleers D. 2000. Adult human pancreatic duct and islet cells exhibit similarities in expression and differences in phosphorylation and complex formation of the homeodomain protein Ipf-1. *Diabetes* **49**:571–579. doi: [10.2337/diabetes.49.4.571](https://doi.org/10.2337/diabetes.49.4.571), PMID: [10871194](https://pubmed.ncbi.nlm.nih.gov/10871194/)
- Heins N**, Englund MC, Sjöblom C, Dahl U, Tønning A, Bergh C, Lindahl A, Hanson C, Semb H. 2004. Derivation, characterization, and differentiation of human embryonic stem cells. *Stem Cells* **22**:367–376. doi: [10.1634/stemcells.22-3-367](https://doi.org/10.1634/stemcells.22-3-367), PMID: [15153613](https://pubmed.ncbi.nlm.nih.gov/15153613/)
- Hoops TC**, Rindler MJ. 1991. Isolation of the cDNA encoding glycoprotein-2 (GP-2), the Major zymogen granule membrane protein. homology to uromodulin/Tamm-Horsfall protein. *The Journal of Biological Chemistry* **266**:4257–4263. PMID: [1999417](https://pubmed.ncbi.nlm.nih.gov/1999417/)
- Hrvatin S**, O'Donnell CW, Deng F, Millman JR, Pagliuca FW, Dilorio P, Rezania A, Gifford DK, Melton DA. 2014. Differentiated human stem cells resemble fetal, not adult,  $\beta$  cells. *PNAS* **111**:3038–3043. doi: [10.1073/pnas.1400709111](https://doi.org/10.1073/pnas.1400709111), PMID: [24516164](https://pubmed.ncbi.nlm.nih.gov/24516164/)
- Jennings RE**, Berry AA, Kirkwood-Wilson R, Roberts NA, Hearn T, Salisbury RJ, Blaylock J, Piper Hanley K, Hanley NA. 2013. Development of the human pancreas from foregut to endocrine commitment. *Diabetes* **62**:3514–3522. doi: [10.2337/db12-1479](https://doi.org/10.2337/db12-1479), PMID: [23630303](https://pubmed.ncbi.nlm.nih.gov/23630303/)
- Jennings RE**, Berry AA, Strutt JP, Gerrard DT, Hanley NA. 2015. Human pancreas development. *Development* **142**:3126–3137. doi: [10.1242/dev.120063](https://doi.org/10.1242/dev.120063), PMID: [26395141](https://pubmed.ncbi.nlm.nih.gov/26395141/)
- Jermendy A**, Toschi E, Aye T, Koh A, Aguayo-Mazzucato C, Sharma A, Weir GC, Sgroi D, Bonner-Weir S. 2011. Rat neonatal beta cells lack the specialised metabolic phenotype of mature beta cells. *Diabetologia* **54**:594–604. doi: [10.1007/s00125-010-2036-x](https://doi.org/10.1007/s00125-010-2036-x), PMID: [21240476](https://pubmed.ncbi.nlm.nih.gov/21240476/)
- Johnson JD**. 2016. The quest to make fully functional human pancreatic beta cells from embryonic stem cells: climbing a mountain in the clouds. *Diabetologia* **59**:2047–2057. doi: [10.1007/s00125-016-4059-4](https://doi.org/10.1007/s00125-016-4059-4), PMID: [27473069](https://pubmed.ncbi.nlm.nih.gov/27473069/)
- Karakikes I**, Ameen M, Termglinchan V, Wu JC. 2015. Human induced pluripotent stem cell-derived cardiomyocytes: insights into molecular, cellular, and functional phenotypes. *Circulation Research* **117**:80–88. doi: [10.1161/CIRCRESAHA.117.305365](https://doi.org/10.1161/CIRCRESAHA.117.305365), PMID: [26089365](https://pubmed.ncbi.nlm.nih.gov/26089365/)
- Kelly OG**, Chan MY, Martinson LA, Kadota K, Ostertag TM, Ross KG, Richardson M, Carpenter MK, D'Amour KA, Kroon E, Moorman M, Baetge EE, Bang AG. 2011. Cell-surface markers for the isolation of pancreatic cell types derived from human embryonic stem cells. *Nature Biotechnology* **29**:750–756. doi: [10.1038/nbt.1931](https://doi.org/10.1038/nbt.1931), PMID: [21804561](https://pubmed.ncbi.nlm.nih.gov/21804561/)
- Liu H**, Yang H, Zhu D, Sui X, Li J, Liang Z, Xu L, Chen Z, Yao A, Zhang L, Zhang X, Yi X, Liu M, Xu S, Zhang W, Lin H, Xie L, Lou J, Zhang Y, Xi J, et al. 2014. Systematically labeling developmental stage-specific genes for the study of pancreatic  $\beta$ -cell differentiation from human embryonic stem cells. *Cell Research* **24**:1181–1200. doi: [10.1038/cr.2014.118](https://doi.org/10.1038/cr.2014.118), PMID: [25190258](https://pubmed.ncbi.nlm.nih.gov/25190258/)
- Manjunatha B**, Nithin MD, Sameer S. 2012. Cross sectional study to determine gestational age by metrical measurements of foot length. *Egyptian Journal of Forensic Sciences* **2**:11–17. doi: [10.1016/j.ejfs.2011.11.003](https://doi.org/10.1016/j.ejfs.2011.11.003)
- Miyatsuka T**, Matsuoka TA, Sasaki S, Kubo F, Shimomura I, Watada H, German MS, Hara M. 2014. Chronological analysis with fluorescent timer reveals unique features of newly generated  $\beta$ -cells. *Diabetes* **63**:3388–3393. doi: [10.2337/db13-1312](https://doi.org/10.2337/db13-1312), PMID: [24834978](https://pubmed.ncbi.nlm.nih.gov/24834978/)
- Nair G**, Hebrok M. 2015. Islet formation in mice and men: lessons for the generation of functional insulin-producing  $\beta$ -cells from human pluripotent stem cells. *Current Opinion in Genetics & Development* **32**:171–180. doi: [10.1016/j.gde.2015.03.004](https://doi.org/10.1016/j.gde.2015.03.004), PMID: [25909383](https://pubmed.ncbi.nlm.nih.gov/25909383/)
- Nelson SB**, Schaffer AE, Sander M. 2007. The transcription factors Nkx6.1 and Nkx6.2 possess equivalent activities in promoting beta-cell fate specification in Pdx1+ pancreatic progenitor cells. *Development* **134**:2491–2500. doi: [10.1242/dev.002691](https://doi.org/10.1242/dev.002691), PMID: [17537793](https://pubmed.ncbi.nlm.nih.gov/17537793/)
- Ohlsson H**, Karlsson K, Edlund T. 1993. IPF1, a homeodomain-containing transactivator of the insulin gene. *The EMBO Journal* **12**:4251–4259. PMID: [7901001](https://pubmed.ncbi.nlm.nih.gov/7901001/)
- Pagliuca FW**, Millman JR, Gürtler M, Segel M, Van Dervort A, Ryu JH, Peterson QP, Greiner D, Melton DA. 2014. Generation of functional human pancreatic  $\beta$  cells in vitro. *Cell* **159**:428–439. doi: [10.1016/j.cell.2014.09.040](https://doi.org/10.1016/j.cell.2014.09.040), PMID: [25303535](https://pubmed.ncbi.nlm.nih.gov/25303535/)
- Pan FC**, Wright C. 2011. Pancreas organogenesis: from bud to plexus to gland. *Developmental Dynamics* **240**:530–565. doi: [10.1002/dvdy.22584](https://doi.org/10.1002/dvdy.22584), PMID: [21337462](https://pubmed.ncbi.nlm.nih.gov/21337462/)
- Playne R**, Connor B. 2017. Understanding Parkinson's Disease through the Use of Cell Reprogramming. *Stem Cell Reviews and Reports* **13**:151–169. doi: [10.1007/s12015-017-9717-5](https://doi.org/10.1007/s12015-017-9717-5), PMID: [28083784](https://pubmed.ncbi.nlm.nih.gov/28083784/)
- Polak M**, Bouchareb-Banaei L, Scharfmann R, Czernichow P. 2000. Early pattern of differentiation in the human pancreas. *Diabetes* **49**:225–232. doi: [10.2337/diabetes.49.2.225](https://doi.org/10.2337/diabetes.49.2.225), PMID: [10868939](https://pubmed.ncbi.nlm.nih.gov/10868939/)
- Rezania A**, Bruin JE, Arora P, Rubin A, Batushansky I, Asadi A, O'Dwyer S, Quiskamp N, Mojibian M, Albrecht T, Yang YH, Johnson JD, Kieffer TJ. 2014. Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nature Biotechnology* **32**:1121–1133. doi: [10.1038/nbt.3033](https://doi.org/10.1038/nbt.3033), PMID: [25211370](https://pubmed.ncbi.nlm.nih.gov/25211370/)



- Riedel MJ**, Asadi A, Wang R, Ao Z, Warnock GL, Kieffer TJ. 2012. Immunohistochemical characterisation of cells co-producing insulin and glucagon in the developing human pancreas. *Diabetologia* **55**:372–381. doi: [10.1007/s00125-011-2344-9](https://doi.org/10.1007/s00125-011-2344-9), PMID: [22038519](https://pubmed.ncbi.nlm.nih.gov/22038519/)
- Russ HA**, Parent AV, Ringler JJ, Hennings TG, Nair GG, Shveygert M, Guo T, Puri S, Haataja L, Cirulli V, Blelloch R, Szot GL, Arvan P, Hebrosk M. 2015. Controlled induction of human pancreatic progenitors produces functional beta-like cells in vitro. *The EMBO Journal* **34**:1759–1772. doi: [10.15252/embj.201591058](https://doi.org/10.15252/embj.201591058), PMID: [25908839](https://pubmed.ncbi.nlm.nih.gov/25908839/)
- Salisbury RJ**, Blaylock J, Berry AA, Jennings RE, De Krijger R, Piper Hanley K, Hanley NA. 2014. The window period of NEUROGENIN3 during human gestation. *Islets* **6**:e954436. doi: [10.4161/19382014.2014.954436](https://doi.org/10.4161/19382014.2014.954436), PMID: [25322831](https://pubmed.ncbi.nlm.nih.gov/25322831/)
- Scharfmann R**, Pechberty S, Hazhouz Y, von Bülow M, Bricout-Neveu E, Grenier-Godard M, Guez F, Rachdi L, Lohmann M, Czernichow P, Ravassard P. 2014. Development of a conditionally immortalized human pancreatic  $\beta$  cell line. *Journal of Clinical Investigation* **124**:2087–2098. doi: [10.1172/JCI72674](https://doi.org/10.1172/JCI72674), PMID: [24667639](https://pubmed.ncbi.nlm.nih.gov/24667639/)
- Scharfmann R**, Rachdi L, Ravassard P. 2013. Concise review: in search of unlimited sources of functional human pancreatic beta cells. *STEM CELLS Translational Medicine* **2**:61–67. doi: [10.5966/sctm.2012-0120](https://doi.org/10.5966/sctm.2012-0120), PMID: [23283495](https://pubmed.ncbi.nlm.nih.gov/23283495/)
- Segerstolpe Å**, Palasantza A, Eliasson P, Andersson EM, Andréasson AC, Sun X, Picelli S, Sabirsh A, Clausen M, Bjursell MK, Smith DM, Kasper M, Ämmälä C, Sandberg R. 2016. Single-Cell transcriptome profiling of human pancreatic islets in Health and type 2 Diabetes. *Cell Metabolism* **24**:593–607. doi: [10.1016/j.cmet.2016.08.020](https://doi.org/10.1016/j.cmet.2016.08.020), PMID: [27667667](https://pubmed.ncbi.nlm.nih.gov/27667667/)
- Semb H**. 2004. Cadherins in development. *Handbook of Experimental Pharmacology*:53–68. doi: [10.1007/978-3-540-68170-0\\_3](https://doi.org/10.1007/978-3-540-68170-0_3), PMID: [20455090](https://pubmed.ncbi.nlm.nih.gov/20455090/)
- Spangrude GJ**, Heimfeld S, Weissman IL. 1988. Purification and characterization of mouse hematopoietic stem cells. *Science* **241**:58–62. doi: [10.1126/science.2898810](https://doi.org/10.1126/science.2898810), PMID: [2898810](https://pubmed.ncbi.nlm.nih.gov/2898810/)
- Subramanian A**, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *PNAS* **102**:15545–15550. doi: [10.1073/pnas.0506580102](https://doi.org/10.1073/pnas.0506580102), PMID: [16199517](https://pubmed.ncbi.nlm.nih.gov/16199517/)
- Sugiyama T**, Rodriguez RT, McLean GW, Kim SK. 2007. Conserved markers of fetal pancreatic epithelium permit prospective isolation of islet progenitor cells by FACS. *PNAS* **104**:175–180. doi: [10.1073/pnas.0609490104](https://doi.org/10.1073/pnas.0609490104), PMID: [17190805](https://pubmed.ncbi.nlm.nih.gov/17190805/)
- Sussel L**, Kalamaras J, Hartigan-O'Connor DJ, Meneses JJ, Pedersen RA, Rubenstein JL, German MS. 1998. Mice lacking the homeodomain transcription factor Nkx2.2 have diabetes due to arrested differentiation of pancreatic beta cells. *Development* **125**:2213–2221. PMID: [9584121](https://pubmed.ncbi.nlm.nih.gov/9584121/)
- Trott J**, Tan EK, Ong S, Titmarsh DM, Denil S, Giam M, Wong CK, Wang J, Shboul M, Eio M, Cooper-White J, Cool SM, Rancati G, Stanton LW, Reversade B, Dunn NR. 2017. Long-Term Culture of Self-renewing pancreatic progenitors derived from human pluripotent stem cells. *Stem Cell Reports* **8**:1675–1688. doi: [10.1016/j.stemcr.2017.05.019](https://doi.org/10.1016/j.stemcr.2017.05.019), PMID: [28591650](https://pubmed.ncbi.nlm.nih.gov/28591650/)
- van de Bunt M**, Lako M, Barrett A, Gloy AL, Hansson M, McCarthy MI, Beer NL, Honoré C. 2016. Insights into islet development and biology through characterization of a human iPSC-derived endocrine pancreas model. *Islets* **8**:83–95. doi: [10.1080/19382014.2016.1182276](https://doi.org/10.1080/19382014.2016.1182276), PMID: [27246810](https://pubmed.ncbi.nlm.nih.gov/27246810/)
- Villasenor A**, Chong DC, Cleaver O. 2008. Biphasic Ngn3 expression in the developing pancreas. *Developmental Dynamics* **237**:3270–3279. doi: [10.1002/dvdy.21740](https://doi.org/10.1002/dvdy.21740), PMID: [18924236](https://pubmed.ncbi.nlm.nih.gov/18924236/)
- Yu S**, Michie SA, Lowe AW. 2004. Absence of the major zymogen granule membrane protein, GP2, does not affect pancreatic morphology or secretion. *Journal of Biological Chemistry* **279**:50274–50279. doi: [10.1074/jbc.M410599200](https://doi.org/10.1074/jbc.M410599200), PMID: [15385539](https://pubmed.ncbi.nlm.nih.gov/15385539/)



## Appendix 1

### Methods supplementary information

#### Determination of the developmental stage

Weeks of development (WD) were determined based on the length and the anatomy of the foot and the hand. We also used the Human Embryo Resource website ([https://embryology.med.unsw.edu.au/embryology/index.php/Embryonic\\_Development](https://embryology.med.unsw.edu.au/embryology/index.php/Embryonic_Development)). Then, the precise age was calculated with the length of the foot and the following formula: Age = 4.11 X Foot Length + 5.60 (*Manjunatha et al., 2012*) from the Egyptian Journal of Forensic Sciences).

#### Replicates

All replicates in this study are biological replicates. The 'n' in the figure legends indicates the numbers of independent biological replicates at each developmental stage. Each flow cytometry plot displays the results from a single human fetal pancreas. An unpaired Student's t-test was performed on the FACS data of **Figures 2D, 5D–E** and **7E**. Below is the number of independent pancreas analyzed:

- Figure 1A: 9 pancreases at 9.7WD
- Figure 1C–D: 3 pancreases at 9WD
- Figure 2A: 4 pancreases at 9.4WD
- Figure 2C: 2 pancreases at 7WD, 9 pancreases at 8.4WD, 4 pancreases at 9.4WD, 5 pancreases at 12WD
- Figure 5B: 3 pancreases at 9.4WD
- Figure 5C: 2 pancreases at 7WD, 3 pancreases at 8.4WD, 3 pancreases at 11.3WD.
- Figure 7B: 3 pancreases at 9.4WD
- Figure 7C: three independent FACS analyses assayed on AD3.1 cell line differentiated until stage 5 day 1.
- Figure 7E–F: three independent FACS analysis assayed on three different hPSC lines differentiated from stage 2 to 5: SA121 hESC, AD2.1 iPSC, AD3.1 iPSC. Figures display the data from the three lines.

The variability between individual fetal pancreases at the same WD is displayed in **Figures 2D, 5D–E** and **7E–F**.

#### qPCR

All qPCR are from three independent experiments derived from the sorting of three independent fetal pancreases.

#### Global transcriptomic analysis

We performed a global transcriptomic analyses on three independent pancreases at 9WD and three at 11WD.

The following populations were sorted:

- At 9WD: GP2<sup>+</sup> (triplicate), GP2<sup>−</sup> (triplicate), E<sup>low</sup> (duplicate) and M (triplicate)
- At 11WD: GP2<sup>hi</sup> (triplicate), GP2<sup>+</sup> (triplicate), GP2<sup>−</sup> (duplicate), E<sup>low</sup> (triplicate) and M (triplicate)

The third replicate for E<sup>low</sup> population at 9WD and GP2<sup>−</sup> population at 11WD were excluded due to the poor quality of the RNA.